Advanced telecommunication technologies for low-delay access to electronic medical records

Chen, Ziyang

Awarding institution:
King's College London

Advanced Telecommunication Technologies for Low-Delay Access to Electronic Medical Records

by

Ziyang Chen

A Thesis Submitted for the Degree of
Doctor of Philosophy
at King's College London

December 2018

To my family, best friends and Mohammad

Acknowledgements

This thesis would not have been possible without the guidance and the support of several individuals.

First of all, I owe my deepest gratitude to my supervisor Prof. Mohammad Shikh-Bahaei for his continuous support, relentless and insightful guidance. Without your help, your motivation, knowledge and long-lasting experience are irreplaceable assets for me. I will never forget the moment you helped me to overcome the difficulties of my research and life. You never hesitate to share great ideas and experience to me, which are invaluable to me. Thank you, sir.

I would like to thank Prof. Paul Luff for his support and ideas of management. His helpful advice and useful guidance related to my research which add an extra dimension to my work. I would also like to thank Dr. Hadi Saki for his kind help and guidance during the last year of my PhD life. I have never seen someone like Hadi, explaining every detail and sharing many research ideas.

I would like to thank Xi Lu and Tamanna Shikh-Bahaei for being my second author in one of my published paper and invaluable work and suggestions in my research life. I am particularly grateful to all my current and erstwhile colleagues at CTR. Thank you guys for all the support and cooperation you have extended to me and for making my stay at CTR as the most memorable time of my life.

Last, but not the least, I wish to thank my parents, who have always stood by me through thick and thin. It is to them that I owe all the success in life.

Abstract

With the advanced telecommunication technologies, the patients can receive a good quality of medical care. In the past decades, many telecommunication technologies have been proposed and used to process patients' electronic medical records (EMR). For the most of current EMR systems, patients only can get access to some simple medical files, for example, the personal details, description of health condition from physicians, chart records between patients and physicians, and payment bill, etc. However, with more detailed medical records, physicians could provide more effective medical services, which is not only saving the vital time for patients but also medical expenses. In this research, we proposed the systems that patients and physicians can get access to the EMR including medical images and videos with low transmission delay at any time. The primary objective of this research is to come up with new allocation systems and cache patient's EMR in different location of the patient's daily life. The patients and physicians can access EMR timely when needed. The proposed systems could decrease the transmission delay and provide more effective and efficient medical care for patients. Three systems have been proposed based on femtocaching, dynamic vision sensor, edge computing, game theory and deep Q-learning in this research.

For the first system, we allocate patient's medical records (text/words, images, and videos) to the femtocachings which are set up close to the nearest hospitals of the home area, workplace, family home, friends' home and other places according to the patient's social life. Femtocaching has been used to choose and cache the best medical files. Dynamic vision sensor has been proposed to decrease the size of medical video files due to the fact that the medical video files account for a large part of EMR. Dynamic vision sensor takes video according to the reflection of light and how fast the subjects move. It is the first time that femtocaching technology and dynamic vision sensor have been proposed in the health care area.

Different from the previous system, for the second system, an auction-based and non-cooperative game theory algorithm based on edge computing has been proposed for patients to share their edge devices to others instead of themselves only. Edge computing has been used to process medical records instead of cloud computing. There has been a large volume of research of cloud computing that has been conducted in the health care area in the last decade. However, there is an increase in the amount of medical data correlating with the rapidly increasing number of internet of things (IoT) devices. Cloud computing, especially, is not efficient enough for those health devices that require very short response times, this could cause intolerable network latency. Edge computing could cache and process computing tasks at the edge of the network without transferring to the cloud network and protect the privacy of patients. A competitive game mechanism is proposed by combining with the patient's health condition and telecommunication channel quality on the edge. The system would compare the priority of patients and allocate medical files based on the results of the game theory algorithm.

In our third system, we proposed a new sharing algorithm based on deep Q-learning to allocate and cache the medical records of multiple patients. The new system provides more fair and humanized medical care services comparing with the last system. In our last system, all the medical parameters in the game theory allocation algorithm have the same weights, which would cause the situation that the patient has higher illness severity but lower priority. However, in this new system, the patients with a higher level of illness severity have higher priority and lower waiting time to be served. The system would not stop seeking the caching path for patients until the best Q value is found.

At last, according to the overall picture of the research conducted, the main conclusions together with some directions for the future works are presented.

Table of Contents

List of Tables

List of Figures

List of Acronyms

AI Artificial Intelligence.

CIF Common International Format.

CNNs Convolutional Neural Networks.

DQN Deep Q-networks.

DVRs Digital Video Recorders.

DVS Dynamic Vison Sensor.

EMI Electromagnetic Interference.

EMR Electronic Medical Records.

FBS Femtocell Base Station.

GP General Practitioner.

GPS Globe Position System.

IoT Internet of Things.

LTE Long Term Evolution.

MRI Magnetic Resonance Imaging.

QoS Quality of Service.

RAN Radio Access Network.

RNNs Recurrent Neural Networks.

SOI Severity of Illness.

SNR Signal-Noise Rate.

UTs User Terminals.

VCG Vickrey-Clarke-Groves.

Wi-Fi Wireless-Fidelity.

Chapter 1

Introduction

Telemedicine is the use of telecommunication and information technologies to provide medical care from a distance [1]. Medical information can be exchanged from one site to another via electronic communications to improve, maintain, or assist patients' health status [2].

In the early days, telemedicine was mainly used to connect doctors working with a patient in one place with medical experts in other places [3]. Telemedicine is very beneficial for people in rural areas where clinical professionals are not available or for patients who are difficult to reach [4]. Throughout the next few decades, the equipment needed for telemedicine is still complex and expensive, and the usage of telemedicine is growing, but it is still very limited. The rise of the internet era has brought with it profound changes in the practice of telemedicine [5]. It not only improves the popularization of smart devices for high-quality medical files transmission, but also provides the possibility of remote medical care for patients in the home, workplace, or other places and an alternative to professional care for on-site medical treatment. Internet-based health care has made phenomenal growth over the past decades and is becoming an increasingly important part of healthcare infrastructure [6].

By applying telecommunication technologies, all patients, physicians, hospitals, and other medical equipment were connected effectively. For the patients, medical gateway, such as mobile devices, receives medical data from patients and sends to physicians, hospital, and emergency department. For the physician, they receive the medical data from telecommunication devices and provide the corresponding medical services, for example inquiring the health condition of patients, and giving suggestions and instructions, and ambulances if the urgent situation happened. According to the research, around 40 percent of doctors reported that they currently access their telemedicine programs through a mobile computer or tablet, the ability to access patient information directly from those same machines will greatly streamline their processes [7]. For the tele-medical care, the satisfaction rates of patients are consistently quite high. More than 80% of patients find their telemedicine care to be satisfactory [8].

Electronic Medical Records (EMR) is an essential part of telemedicine, which can improve the quality of patient care and decrease medical errors [9]. In the early days, most patient care records are paper-based and not well-organized. They are usually incomplete and can hardly be accessed in time [10], which means the patient cannot get efficient and effective medical support. An important trend in medical informatics is the adoption of electronic patient record systems that facilitate access to clinical information and work toward preventing the loss or misplacement of information [11] [12].

With telecommunication technologies, patient medical records can be well-organized and kept more secure [13]. Medical records can be used without considering distance barriers, which permits communications between patient and medical staff with both convenience and accuracy. It is also used to save lives in critical care and emergency situation [14]. According to the research [15], 28 percent of physicians who use telemedicine services conduct their video visits through an EMR platform, with an additional 38 percent reporting that they would like to have that capability in the future.

In this research, novel and advanced telecommunication technologies including femtocaching technology, edge computing, dynamic vision sensor, auction-based non-cooperative game theory, and deep Q-learning have been proposed to access medical records with low transmission delay and save more time for clinical professions and patients.

1.1 Thesis Contributions

In the past decades, many telecommunication technologies have been proposed and used to process patients' electronic medical records. However, for the most of current EMR systems, patients only can access some simple medical files [16]. In this research, we proposed the systems that patients and physicians can access the whole EMR including medical images and videos with low transmission delay at any time. The primary objective of this research is to allocate and cache patient's EMR in different location of patient's daily life. The patients and physicians can access EMR timely when needed. The proposed systems could decrease the transmission delay and provide more effective and efficient medical care for patients.

This study provides new ideas to current heath care filed combining with the telecommunication viewpoints. More specifically, this study builds three new systems to access and cache medical records with low delay for patients and physicians. New allocation and caching algorithm of medical records have been proposed as well. The results show that the proposed systems could save more time for patients and provide more efferent medical care services comparing with the traditional medical services.

1.1.1 Key Outcomes

The contributions of this thesis cover the simulation and performance analysis of five different scenarios, including the system with location-aware femtocaching, edge computing, DVS camera, auction-based non-cooperative game theory and deep Q-learning. The Key outcomes of this work in the form raised issues and novel solutions, systems design, allocation and caching algorithms, system simulation analysis are summarized as follows.

1. The first system is focused on a new electronic medical records allocation method in order to improve the transmission of medical files over wireless

networks by using location-aware femtocaching and DVS camera. In the previous study [17] [18], femtocaching is proposed to receive and cache popular videos of new media platform. For the users who want to watch the videos, they can download from femtocaching instead of macro base station, thus decreasing the burden of networks. For the dynamic vison sensor, which has been proposed to decrease the size of medical video files. DVS is a very new and novel technology as well and not yet universal, which takes video according to the reflection of light and how fast the subjects move. In the previous research [19] [20] [21] [22], there are a few researches applied DVS into transportation monitoring and pedestrian walking monitoring. DVS camera could save more resources, such as devices battery life and storage capacity of camera to current monitoring system, especially for the patients with sleep disorder. In our proposed system, we divided electronic medical records into three levels according to the type and the size of medical files (words/text, images, and videos). The EMR is allocated in femtocachings efficiently and effectively according to our proposed medical records allocation algorithm by applying knapsack model with penalty minimization. Due to the limited storage capacity of caches, not all the caches have the enough space for medical files, especially medical images and videos. The proposed medical records allocation algorithm would choose the most suitable files to cache based on three important factors. In this system, a new medical video capture method is proposed as well to reduce the size of visual medical files by applying DVS camera technology. We compared the scenarios of system with femtocaching and DVS technology, femtocaching only, and without the proposed system. The simulation results show that the transmission delay is 9.79min, 16.59min, and 124.4min respectively in the best situation. For the worst situation, the simulation results show that the transmission delay is 27.37min, 139.65min, and 247.47min respectively. Therefore, the transmission delay can be improved by 88.94% to 92.13% by applying the system with femtocaching and DVS.

2. In the second system, a sharing system has been proposed for patients to share their edge devices to others nearby. A non-cooperated auction-based game theory mechanism has been applied to compete storage capacity among the patients based on edge computing. Edge computing is proposed to process and cache EMR instead of cloud computing. There is a large volume of researches of cloud computing have been conducted in health care area in last decade [23] [24]. However, there is an increase in the amount of data correlating with the rapidly increasing number of internet of things (IoT) devices. The speed of data transportation is becoming the bottleneck for the cloud-based computing paradigm [25]. Cloud computing, especially, is not efficient enough for those health devices that require very short response times, thus could cause intolerable network latency. Edge computing could cache and process computing task at the edge of the network without transferring to the cloud network and protect the privacy of patients. Edge computing could off load partly or whole data processing from cloud and provide more efficient and safer services. Patient can control their own medical data, made by wearable sensors, at the edge. In the previous research, game theory has been used in many filed, especially, for the network resources and spectrum allocation in telecommunication [26] [27] [28]. In our proposed study, game theory is applied on the edge to compete the allocation and caching priority by combining with patient's health condition and telecommunication channel condition. Four medical factors have been considered, including the severity of illness, the size of the medical records that patient wants to cache, the queueing time that patients have been waited, and the rewards that patients got by sharing edge caching with other patients. In the proposed non-cooperative game, all the patients are players. They use their combined value as a bid to compete.

Nash Equilibrium has been presented as a desirable outcome. The results show that the proposed system can decrease the transmission delay by 56.87% to 93.69% in the best situation and by 25.94% to 57.75% in the worst situation comparing with the system without sharing edge devices. On the other hand, the results show that more shared edge devices, higher channel capacity patients can get, which means the proposed system encourages more patients to share their storage capacity of edge devices to others.

3. In the last system, we considered that all the medical factors have same weight. However, the priority is based on a combination value. It is possible that for the patients who has higher illness servility but cannot get timely and effective health care services. Therefore, in this system, we introduced a new allocation algorithm based on deep Q-learning [29] [30] for patients to cache their medical files to adjacent edge devices. M/M/1 queueing model with preemptive priority is proposed for queueing time of patients. On the other hand, the medical files of patients only can be cached in one edge caching in the last system. However, in this new system, the medical files can be cached in multiple edge caching devices dispersedly. The processing time including queueing time and caching time would be decreased significantly. According to the simulation results, patients could get more efficient medical care by applying the proposed system and save at least 71.14% transmission time comparing with the traditional method. We compared this research with our last system as well, the results show that new allocation algorithm saves 20.57% time more for the whole of system, which means other patients who want to cache medical records in edge devices could get medical care in a quicker manner. The proposed system provides more fair and humanized medical care services, the patients with higher level of illness severity have higher priority and lower waiting time to be served.

1.1.2 List of Publications

The list of publications related to the main contributions of this thesis are stated as follows.

(1) Z. Chen and M. Shikh-Bahaei, "Location-Aware Distributed File Allocation for Low-Delay Access to Electronic Medical Records", *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* ***IEEE****,* pp. 5372-5375, 2016.

(2) Z. Chen, T. Shikh-Bahaei, P. Luff, and M. Shikh-Bahaei, "Edge Caching and Dynamic Vision Sensing for Low Delay Access to Visual Medical Information," *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* ***IEEE****,* pp. 1238-1241, 2017.

(3) Z. Chen, T. Shikh-Bahaei, P. Luff, M. Shikh-Bahaei, "Edge Computing for Low Delay to Cache and Transfer Electronic Medical Records," *Advances in Intelligent System and Computing,* ***Springer****,* vol. 869, pp. 474-486, 2018.

1.2 Outline of the Thesis

The rest of the thesis is organised as follows. Chapter 2 provides the preliminaries on the required background for understanding the research area addressed in this thesis. In addition, this chapter presents the related works in health care based on femtocaching, DVS camera, edge computing, game theory and deep learning.

The main contributions of this thesis, which are related to the application of three distinct area: location-aware femtocaching technology with dynamic vision sensor camera, auction-based non-cooperated game theory based on edge computing and deep Q-learning are discussed in Chapters 3, 4, and 5. Each contribution chapter addresses a new medical files transmission system and scenario. The performance analysis of system and simulation results are presented in the end of each chapter. Based on the overall picture of the research conducted in the thesis, the main conclusions together with some directions for future work are presented in Chapter 6.

Chapter 2

Preliminaries and Related work

This chapter provides detailed overview of the existing research works and telecommunication technologies in health care area. Telecommunication technologies are providing advanced method to traditional medical services. It streamlines the health system resources and brings more applications to give effective medical care. Therefore, a large volume of research with respect to novel medical systems, medical data processing and patient care solutions has been done in order to bring the best benefit for patients or improve the medical care services. At the beginning of this chapter, the mainstream telecommunication technologies in health care area, including wearable sensors, big data and internet of things (IoT), have been addressed. The existing research work is explained. Then, the background studies about electronic medical records and current medical records system and platform are presented, respectively. Last but not least, the related works of the femtocaching, DVS camera, edge computing, game theory and deep learning in health care filed are conducted. It includes the basic background of above technologies and existing research achievements.

2.1 Telecommunication in Health Care Area

Telecommunication brings massive change to current health care area， which provides more efficient medical service to patients. According to the research, it is estimated that the 20 to 64 years old group will decrease to 55% and the over 65 group will increase to 28% by the year 2050, which means more medical services and budget are required [31]. As shown as in Figure 2.1, the expenditures of health care increase with age population. Therefore, many researchers are looking for more efficient method to provide more efficient healthcare.

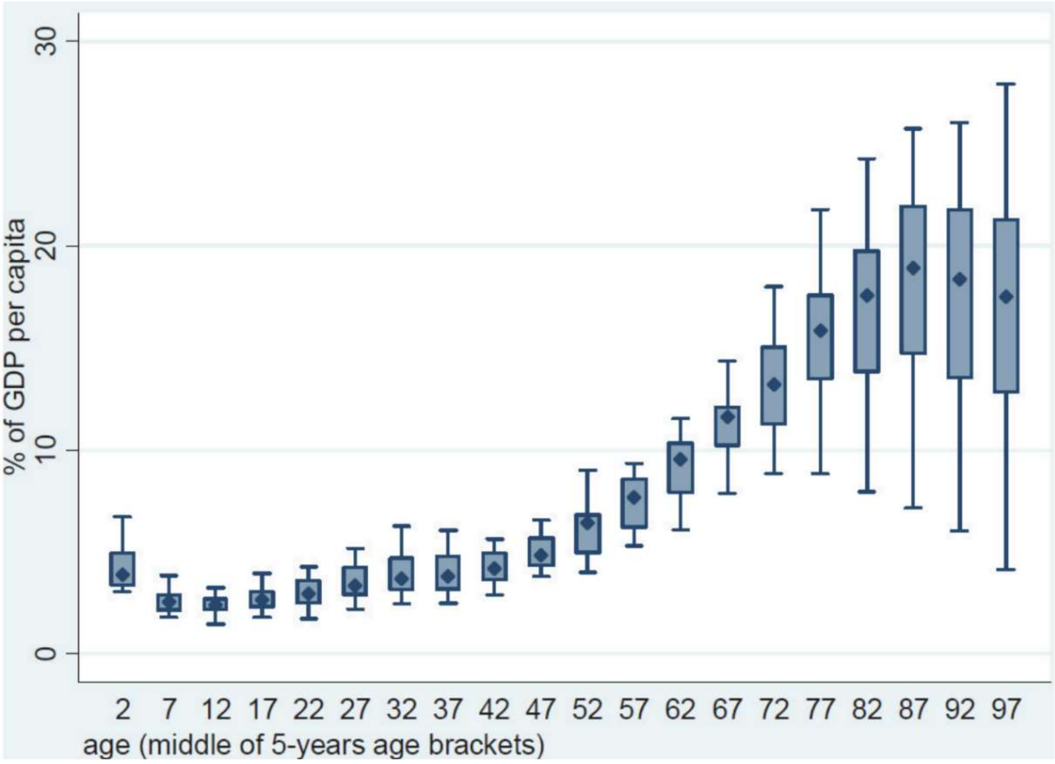

Figure 2.1: Public Heath Care Expenditure by Age Groups [31].

Nowadays, more and more patients are using medical sensors to improve their self-care ability and to obtain more accurate and effective medical services. In the near future, real-time medical data generated by medical sensors will increase significantly [32]. It is estimated that by 2020, the number of internet-connected devices might reach 50 billion, and the increase of network traffic of these devices might have a significant impact on networks, data centers and smart connected products [33]. Wearable sensors have diagnostic, as well as monitoring applications. Physiological measures of interest in rehabilitation include heart rate, respiratory rate, blood pressure, blood oxygen saturation, and muscle activity [34].

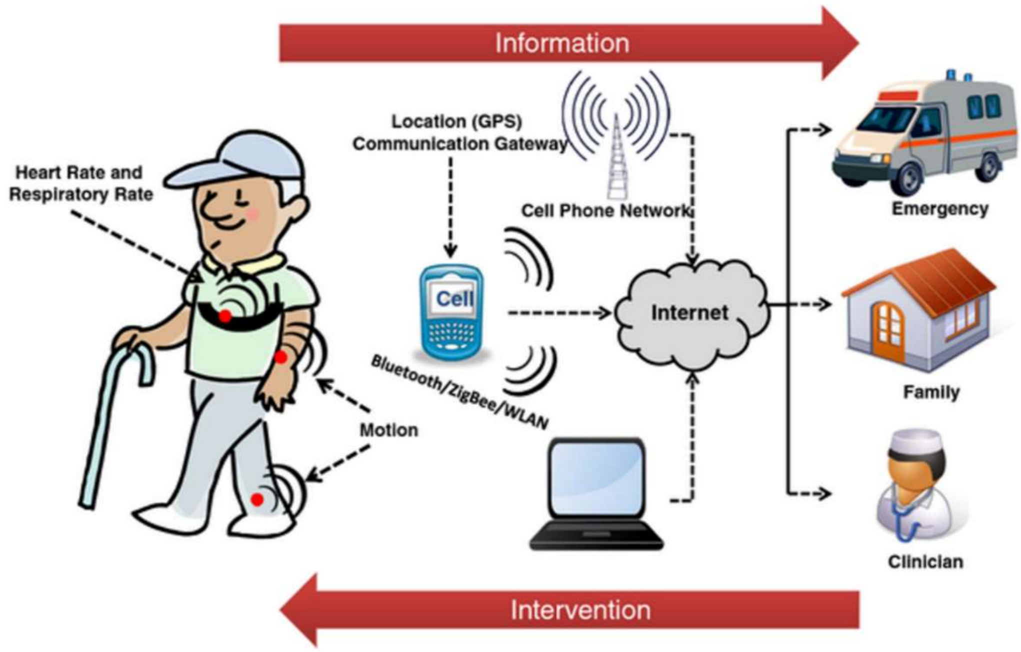

Figure 2.2: Illustration of a Remote Health Monitoring System Based on Wearable Sensors [34].

Wearable sensors are used to collect physiological and movement data from the patient to enable monitoring of the patient's health. Medical sensors are deployed based on the patient's health condition and the patient's consent. For example, when monitoring a patient with congestive heart failure or a patient with chronic obstructive pulmonary disease, the sensors for monitoring vital signs such as heart rate and respiratory rate will be deployed. For rehabilitation monitoring of stroke survivors at home or for monitoring of elder patients with mobile assist devices, wearable sensors will be deployed to collect patient's movement data. Wireless communication is relied upon to transmit patient's data to a mobile phone or an access point and relay the information to a remote center via the internet. Emergency situations (such as falls) are detected by data processing implemented throughout the health care system, and an alert message is sent to the emergency service center to provide immediate assistance to the patient. Family members and caregivers will receive an alert when the patient is under an emergency situation. They can also be notified when a patient needs help, such as taking his/her medication. On the other hand, clinical personnel also can remotely monitor patient's health status and be alerted in case a medical decision has to be made.

In [35], a small and sufficient breathing sensor has been proposed to monitor breathing every day. The sensor is applied to make sounds and get the attention of others when an emergency occurs. It is possible to reduce the number of deaths by allowing faster medical intervention. The breathing detection system aims to use a small, comfortable sensor that may be operated for long periods of time with a small battery. The algorithm is tested in the presence of noise sources and shows an average success rate of 91.3%. In [36], authors present an automatic fall detection system called SmartFall by applying a fall detection sensor. Fall-induced injury has become a leading cause of death for the elderly. Many elderly people rely on canes as an assistive device to overcome problems such as balance disorder and leg weakness, which are believed to have led to many incidents of falling. In the research, authors present the design and the implementation of SmartFall, an automatic fall detection system. The results show that SmartFall system achieves a near perfect fall detection rate in the experiments. In [37], authors present a novel ear-worn sensor to measure heart rate on a mobile patient. The sensor is designed to be physically comfortable with low power consumption and have the ability to manage harsh sensing environments such as those introduced by electromagnetic interference (EMI) and increased temperature or humidity. The experimental results have shown that the sensor system is capable of detecting heart rate at a number of locations on the body.

The data extracted from these medical sensors can provide a huge diagnostic value to the patient's health indicators. However, the number of sensors deployed around the world is growing at a rapid pace and continuously generate enormous amounts of data [38] [39]. Ultimately, these sensors will generate big data [40]. The data we collect may not have any value unless we analyse, interpret, and understand it. To analyse the huge amount of data from various sensors, big data technology has been proposed to address this issue.

In [31]. The author proposed a new medical application for elder people with chronic disease to improve their self-care ability and quality of life by using big data technology. The technology platform manages and processes huge volume data from medical sensors. A data capture system has been proposed to acquire and transfer patients' vital information to the system for processing. In [41], a robust, easy to use health big data system is proposed to collect diabetes patient's medical data and manage diseases. The system also provides a platform for patients to discuss and share treatment management questions, experiences and challenges. The system allows patients to find multiple similar patients to discuss, not only one. The system would send up-to-date information, videos and news articles as well to all the patient to meet all of their disease management needs.

On the other hand, the IoT has been proposed to address the data issues. IoT offers a world of networked devices, cloud-based applications and services, with diverse cooperation mechanisms and where big data analytics enables getting objective data and information, and well-grounded decision making [33]. It seamlessly connects smart sensors attached to the human body for physiological monitoring and daily medication management.

In [42], An IoT-based intelligent home-centric healthcare platform iHome Health-IoT System has been proposed. The network architecture of iHome consists of three network layers: the first one is smart medical service layer. Doctors can efficiently manage a large group of patients from this layer. The second one is medical resource management layer, which includes administration and management of medical resources; the last one is sensor data collecting layer, which is the basis of the entire network. The proposed platform seamlessly fuses IoT devices (e.g., wearable sensors and intelligent medicine packages) with in-home healthcare services (e.g., telemedicine) for an improved user experience and service efficiency. The feasibility of the implemented iHome Health-IoT platform has been proven in field trials. In [43], author proposed a novel architectural framework called H3IoT to monitor health of elderly people at home based on IoT technology. A 5-layered system was proposed including user application layer, internet application layer, information processing layer, local communication layer and physiological sensing layer for elderly people can be monitored at home by their relatives, doctors, nearby hospitals, and care takers where in staying at remote location. H3IoT has many advantages as it is mobility, cheap, easy to use, simple layered design, and delay tolerant.

2.2 Electronic Medical Records

Electronic Medical Records integrate patients' clinical data and constitute a main source of reference for their care [11]. It is very important for the hospital or general practitioner (GP) to access patient's EMR in order to provide efficient, effective, and safe medical services [44] [45]. In early times, like some still in use today, patient's medical records are based on pen and paper. These paper-based medical records are not well organised, which are usually incomplete and can hardly be accessed in time [10], which means the patient cannot get timely and effective medical support. More inherent problems with paper records include legibility, clarity and the ease with which they can be lost. Multiple records must be created and maintained, such as a medical chart and an X-ray folder. Each usually requires separate storage and maintenance. This gets even more complicated when patients visit multiple facilities or departments [46].

In order to provide efficient and precise medical care [2], an important trend in medical informatics is the adoption of electronic patient record systems that facilitate access to clinical information and work toward preventing the loss or misplacement of information [11]. By implementing EMR, patients' medical data can be tracked over an extended period of time by multiple healthcare providers [47]. It can help to identify those who are due for preventive check-up and screening and monitor how each patient measures up to certain requirements like vaccinations and blood pressure readings. On the other side, EMR is universal, which means that instead of having different charts at different healthcare facilities [48], a patient will have one electronic chart that can be accessed from any healthcare facility using EMR software [49].

In the past decades, many telecommunication technologies have been proposed and used to process patients' EMR [14]. Many EMR systems have been designed and developed to process medical data. These technologies and systems permit communications between patient and medical staff with both convenience and fidelity, as well as the transmission of medical, imaging and health informatics data from one site to another. Doctors can quickly log in to their protected accounts from any computer or mobile device to pull up patient information for a remote appointment. The online patient portals allow patients to access their own records without making a trip to the office, which is especially important for those who receive their care virtually [50]. On the other side, health care providers or organisations do not provide effective access for patients to their own data and they show little willingness to share patient's medical data to other health care providers. However, EMR system is able to exchange cached data among the health care providers according to standard in securing a consensus. Hence having the right EMR system in place is essential for maximizing the quality of medical services.

With the advanced telecommunication technologies, the patient medical records can be accessed efficiently and shared among hospitals, which is essential in delivering the quality of medical care services [13]. The most obvious improvement by using EMR systems is the increase in accuracy and efficiency [46].

2.3 Related Works

The related works of the femtocaching, DVS camera, edge computing, game theory and deep learning in health care filed are conducted in this section. It includes the basic background of above technologies and existing research achievements.

2.3.1 Femtocaching Networks

Femtocaching is femtocell-like base station with weak backhaul links but large storage capacity. Femtocaching was first proposed in [18]. Authors suggest femtocaching network to handle increasing demands for video content in wireless/mobile devices. Femtocaches form a wireless distributed caching network that assists the macro base station by handling requests of popular files that have been cached. Due to the short distances between femtocaches and users, the transmission of cached files can be done very efficiently. Femtocaches can cache popular files and serve requests from mobile user terminals (UTs) by enabling localized communication and hence frequency reuse. Because the most popular files are stored in the cache and are thus always available locally to the UTs that are requesting it. The results show that the number of users that can be served is increased by as much as 400% to 500%.

In [17], authors present a different architecture for the same issues by considering using the wireless terminals themselves as caches, which can distribute video through device-to-device communications. In the system, femtocaching is operated in conjunction with a traditional, macro cellular base station. Clearly, the smaller the percentage of file requests that has to be fulfilled by the macro base station, however, the larger the number of UTs that can be served. The simulation results show that this approach improves video throughput by one to two orders-of-magnitude and allows an improvement in the video transmission delay without deployment of any additional infrastructure.

In [51], author proposed a new content updating method for femtocaching based on previous researches. The previous work does not provide any method to update contents in caches of helpers when the ranking changes. The number of contents that can be cached is very limited because the sizes of contents are usually huge. If the femtocachings cache unnecessary contents, the maximum benefit from caching might not be achieved.

In [18], femtocaching placement algorithm is proposed to identify the popular contents in the system and cache the popular contents only. The complexity of the proposed method is less than the previous algorithm in femtocaching. The simulation results show that the proposed method has an acceptable performance considering overall cache hit rates and more suitable for updating than the femtocaching placement algorithm.

2.3.2 Dynamic Vision Sensor Camera Technology

A dynamic vision sensor is a sensor that detects temporal contrast of brightness and has the fastest response time compared to conventional frame-based sensors, which detects static brightness per every frame. The fastest response time allows fast motion recognition, which is a very attractive function in the view of consumers. A conventional video camera records at something like 30 or perhaps even 60 frames per second. However, DVS is event-based. When subjects move faster, DVS takes more frames per second. When subjects move slower, DVS takes less frames per second. In particular, it is low power consumption due to the event-based processing, which is a key feature for mobile applications [52].

In [20], a DVS camera is applied into a traffic intersection, which is shown in Figure 2.3. The sensor outputs a measurement only when the light changes. When the light level crosses that threshold, the sensor produces a “change-detection event” and triggers an illuminance measurement. The results for the entire pixel array can be viewed for a given time slice, but only changing pixels contribute to these images: the positive and negative level-crossing events (white and black, respectively, middle right) and associated illuminance measurements (nonblack areas, bottom right).

In [19], DVS camera is proposed for a person tracking application, which is shown in Figure 2.4. The left two images are from a 2 second video sequence, which have been extracted for illustration. On the right side, the tracking results is provided according to the system simulation. The circles indicate the locations of different persons. A unique ID number identifies the person and an arrow indicates the direction and speed of movement. As we can find from the tow right pictures, ID 198, 155, 171, and 222 have been correctly tracked while ID 227 was a shadow effect that disappeared in the next sequence. The results show that the tracking algorithm could significantly reduce the computational burden comparing to the traditional video-based tracking surveillance systems.

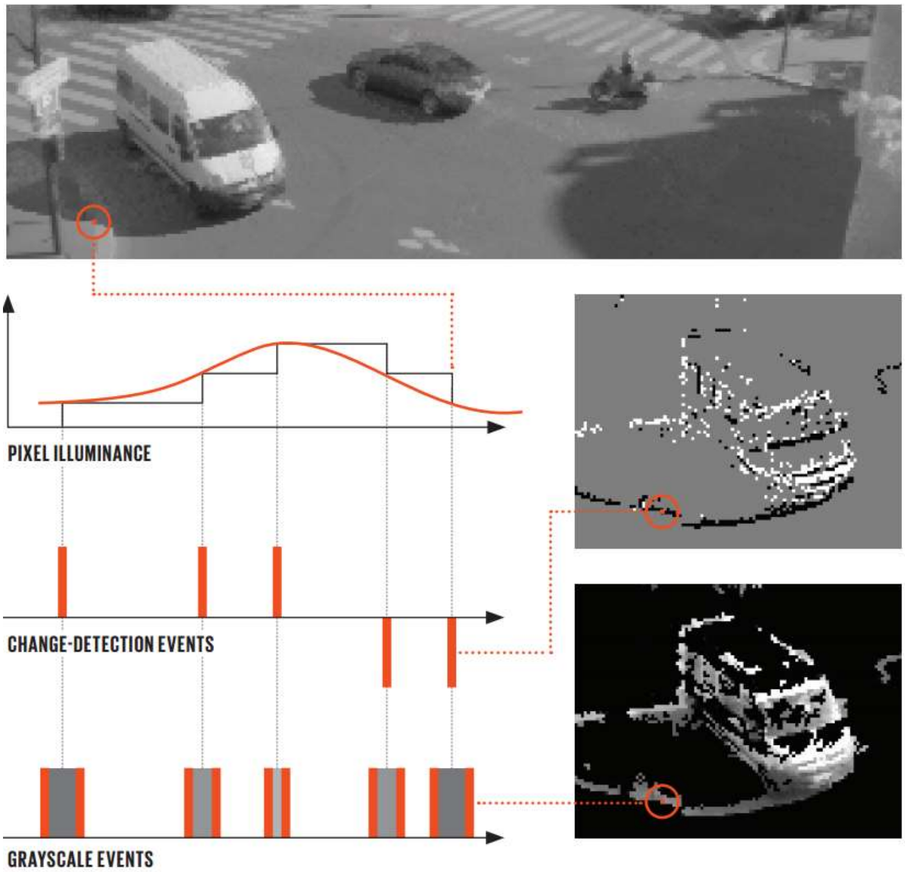

Figure 2.3: A Traffic Intersection by Applying DVS camera [20].

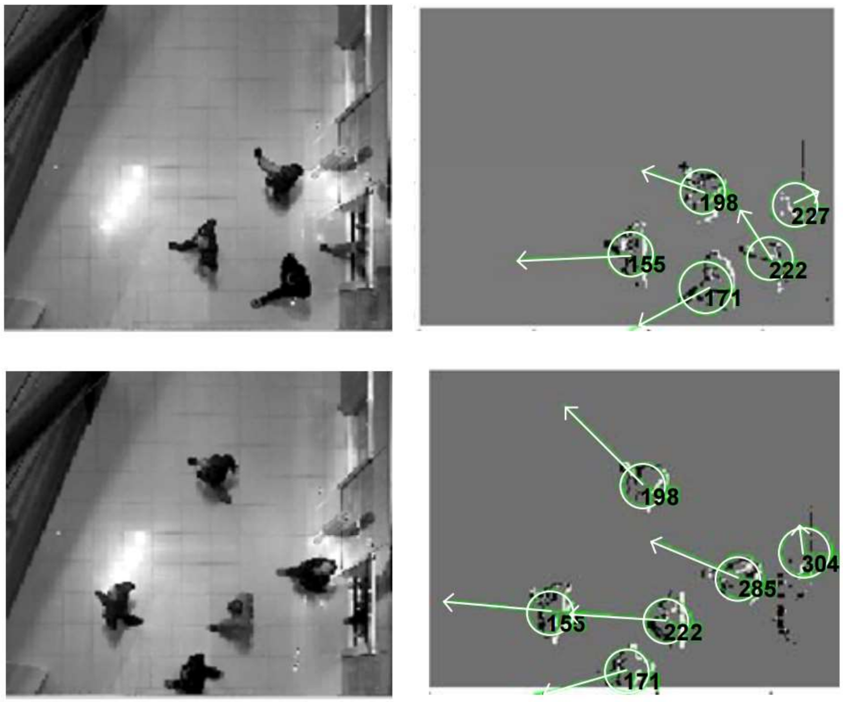

Figure 2.4: A Passenger Tracking Application by Applying DVS Camera [19].

2.3.3 Edge Computing Networks

In the edge computing network, data is processed by the device itself or by a local computer or server, rather than being transmitted to a public cloud or data center platform, which increases the edge responsibility and allows computation and services to be hosted at the edge, and reduces the network latency and bandwidth consumption [53]. Edge computing concept has been discussed only from a theoretical perspective so far. However, there are some related approaches that are similar to this concept [54].

A new edge computing network named Cloudlet has been proposed to overcome long latency of data exchange between the public clouds and edge device through the internet when offloading computation to the public. Cloudlet-based offloading is proposed where mobile devices offload computational to the less resourceful server near the user proximity accessible using Wi-Fi access point [55].

Cloudlet networks can provide cloud computing services to mobile devices, such as smartphones, tablets and wearable devices, within close geographical proximity. The results show that Cloudlet networks could improve the response time of applications running on mobile devices, decrease the latency and provide better resource allocation [56] [57] [58].

In [59], authors proposed FemtoClouds system that provides a dynamic and self-configuring multiple device cloud system to scale the computation of Cloudlet by coordinating multiple mobile devices. FemtoClouds could decrease the network latency during computational offloading to the traditional cloud data center by applying the nearby unutilized mobile devices as edge computing devices. The results show that FemtoClouds system provides a community-based computing service by minimizing the dependent on corporate infrastructure and scaling the computation capability.

In [60], an edge cloud architecture named REPLISOM, has been proposed to reduce cloud responsiveness under long term evolution (LTE) environment when multiple IoT devices replicate memory objects to the edge network. In REPLISON system, instead of pushing the new updated memory object, the edge cloud pulls the memory replica to the respective clone virtual machine. The results show that REPLISOM could decrease the latency and cost during offloading when multiple IoT devices replicate updated object to the nearby edge cloud. The analysis and numerical evaluation show that the proposed system could reduce the cost for cloud offloading, transmission delay, and energy consumption significantly in terms of the IoT applications given the massive number of devices but with tiny memory sizes.

2.3.4 Game Theory

Game theory is a mathematical tool used in understanding and modelling competitive situations which imply the interaction of rational decision makers with mutual and possibly conflicting interests. Nowadays game theory has been widely applied to different subjects. In telecommunication areas, game theory has been proposed to efficiently allocate the resources of wireless networks to users, optimize the performance of networks and control distributed systems [61].

In [28], authors propose a novel method to allocate wireless resources based on game theory in terms of transmit power and wireless spectrum. In other to find the optimal allocation results, each mobile user would bid for the limited wireless resources from physical substrate wireless networks and compete the same resources with others. Three different resources request strategies are proposed including price-based, correlation-based, and water-filling-based respectively. The simulated results show the benefits and revenues of the virtual wireless networks and physical substrate wireless network are both optimized, thus maximizing the aggregate spectrum efficiencies over wireless-resources-virtualized mobile cloud-computing wireless networks.

In [26], it presents the current state of optimizing resource allocation in LTE/LTE-A networks based on game theory. In current wireless network, resource allocation is becoming more complicated in the results of multi-cell, multi-user, multi-antenna and multipath fading. By applying the game theory, the results show the system could achieve better performances in term of system delay, network throughput, fairness index, and packet loss ratio, etc.

2.3.5 Deep Learning Algorithm

Deep learning allows computational and processing systems that are composed of different multiple layers based on neural networks to learn representations of data with multiple levels of abstraction [62], which is different from traditional machine learning that how to extract information from the raw data [63]. More recently, deep learning has been applied to process aggregated electronic medical records in health care area, including both structured (e.g. laboratory tests, medications and diagnosis, etc.) and unstructured (e.g. free-text clinical notes) data.

The results from current research show that deep learning obtains better results than conventional machine learning model. The deep learning applied to the health care has been mostly based on two neural networks, including convolutional neural networks (CNNs) [64] and recurrent neural networks (RNNs) [65].

In [66], authors proposed a four-layer convolutional neural networks based on deep learning to predict congestive heart failure and chronic obstructive pulmonary disease. The results show that proposed networks could bring significant advantages over the baselines. In [67], an end-to-end deep learning dynamic network named DeepCare based on recurrent neural networks has been proposed to infer current illness states and predicts future medical outcomes. Moreover, the medical interventions are combined into the model to dynamically shape the predictions. DeepCare was evaluated for disease progression modelling intervention recommendation and future risk prediction on diabetes and mental health patients. In [68], an end-to-end recurrent neural networks model called Doctor AI has been proposed to predict diagnoses according to the health care history of patients. The evaluation showed significantly higher recall than conventional medical services. Patients could get more timely medical intervention and save medical expenses.

Chapter 3

A Location-Aware Distributed File Allocation Method for Low-Delay Access to Electronic Medical Records

3.1 Introduction

Electronic medical records integrate patients' clinical data and constitute a main source of reference for their care [69]. EMR can keep medical files well organized and more secure. It is important for the hospitals or physicians to access patient's EMR in order to provide efficient, effective, and safe health care services [70] [71]. It is important to be able to share patients' medical records among the hospitals as well, which not only provides a vital and efficient reference to patients and physicians, but also saves more medical care expense [72]. However, many health care providers or organisations do not provide effective access for patients to their own data and do not share patient's medical data with other health care providers. One of the important reasons is different hospitals have different medical standards. The medical records from different medical service providers are difficult to share with others. On the other hand, for most of current EMR systems, patients only can access some simple medical files [16]. However, with more detailed medical records, physicians could provide more precise medical services [9]. With the advanced telecommunication technologies, the patient medical records can be accessed timely and shared among hospitals, which is essential in delivering the quality of medical care services [13].

In this chapter, a new medical files allocation system is proposed to transfer, allocate, and cache EMR by applying location-aware femtocaching technology and dynamic vision sensor camera. The purpose of this new system is to provide whole EMR access with low-delay transmission instead of basic medical files. By applying the new system, medical professions can provide effective medical services to patients based on their locations at any time. We applied location-aware femtocaching technology to the nearest hospitals of patient's daily location [73]. Medical records are allocated and cached in different femtocaches by using the proposed allocation algorithm. Femtocaching is femtocell-like base station with large storage capacity. The surest way to increase the system capacity of a wireless link by getting the transmitter and receiver closer to each other, which creates the dual benefits of higher-quality links and more spatial reuse. One of the means to achieve this with higher system capacity and lower delay is based on femtocaching. Femtocaching is proposed to set up near the hospitals in this research. Due to the short distances between requesting devices and femtocaching, the cached medical files can be transmitted efficiently [17]. Femtocaching can specially decrease packet delay when users repeat their requests for the same file.

In particular, for the first time, exploiting the recently developed dynamic vision sensor technology is investigated here to achieve smaller medical visual data files in our proposed tele-health system. According to the size and type of medical files, the whole EMR is divided into three categories (text/word, digital image, and video/audio) in the proposed system. However, the medical video size (e.g. in-patient monitoring), accounts for a big part of the EMR (around 70%), which could cause a huge delay when physicians access to patient medical records. Dynamic vision sensor camera is proposed to decrease the medical video size, which not only saves more time for transmission of visual medical data, but also provides more storage capacity for caching more patients' medical files. The transmission delay is decreased significantly by combing location-aware femtocaching technology and dynamic vision sensor camera.

3.1.1 The Overview of Femtocells Networks

In the mobile wireless networks, the demand for higher data rates and lower power consumptions is continuously increasing, while the capacity provided by the existing macrocell networks is limited [74]. This phenomenon motivates the research and development for femtocell networks, where each femtocell base station (FBS) is installed at each customer's home or office. The FBS is a short-range, low-cost, and low-power base station. It can communicate with macrocell networks by broadband connections such as digital subscriber line (DSL), cable modem, or a separate wireless backhaul channel [75]. Femtocells can provide high data rates and quality of service (QoS) with low transmission power for consumers.

The structure of femtocell networks is shown in Figure 3.1. There are several femto-networks operating within the coverage of the Macro-network. Each femto-network is composed of a base station (abbreviated as femto-BS) and multiple mobile stations (abbreviated as femto-MSs) which is connected to the femto-BS. The macro-network is composed of a base station (abbreviated as Macro-BS) and multiple mobile stations (abbreviated as Macro-MSs) connected to the Macro-BS. Deploying femtocell networks embedded in the macro-cell coverage greatly benefits communication quality in variety of manners [76]. The advantages of applying femtocell are the followings.

- Better coverage and capacity. Due to their short transmit receive distance, femtocells can greatly lower transmit power and achieve a higher signal-to-noise ratio (SNR).
- Improved macro-cell reliability. The network traffic can be absorbed into the femtocell networks. The macro-cell base station can provide more resources and better services to other users.
- Cost benefits. Femtocell deployments will reduce the operating and capital expenditure costs for operators. The deployment of femtocells will reduce the need for adding macro base station towers.

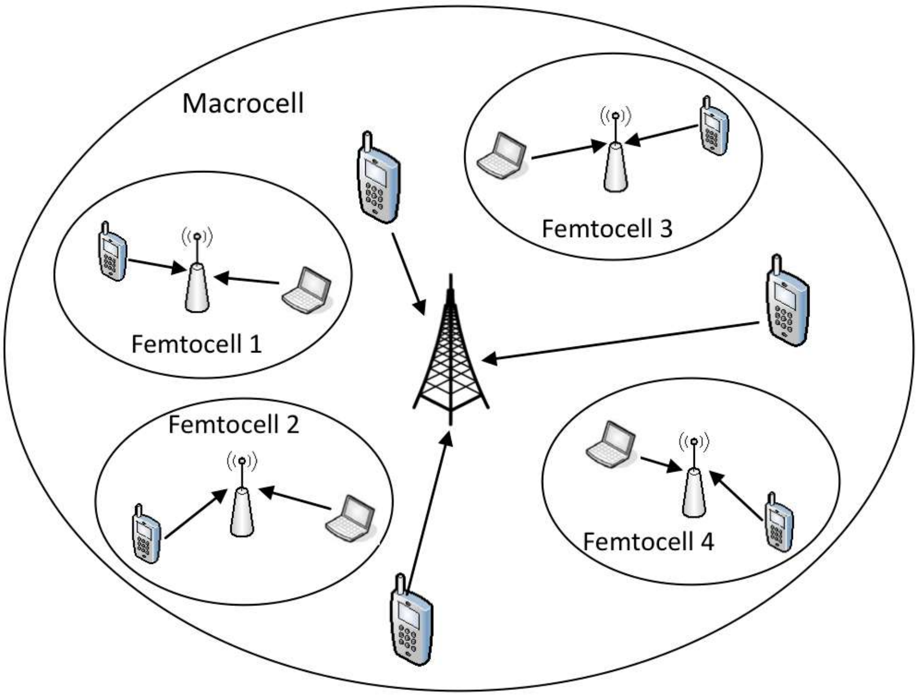

Figure 3.1: The Structure of Femtocell Networks [71].

Femtocaching is base station with large storage capacity based on femtocell network. In [18], author proposed a new wireless distributed caching system to assist the macro base station to handle requests of popular files that have been cached by using the femtocaching. If the user requests a file that is cached in adjacent or local femtocachings, the femtocachings handle the request. The macro base station manages the requests that cannot be handled by femtocaching. Duo to the short distances between femtocachings and users, the transmission of cached files can be done very efficiently.

In this research, we introduced a new method for decreasing the transmission delay of medical records delivery networks. The key idea by applying a distributed femtocaching is to allocate and store the most suitable files, and transmit them, upon request, via short-range wireless links to the nearest hospital. The femtocaching is low-cost because storage capacity has become exceptionally cheap (according to recent prices, two terabytes cost approximately 100 dollars), while the loading of the files to the femtocaching can occur through a low-rate (and thus cheap and robust) backhaul links at low demand times, such as, at night [18].

3.1.2 The Overview of Location-aware Network

Location-aware technology is any technology that is able to detect its current location and then manipulate this data to control events and information. Location is detected through the use of various sensors and methods of calculating geographical location such as through globe position system (GPS) technology and cell tower triangulation. Location-aware devices are able to give directions relative to the location, access geographically aware services and data, as well as broadcast the location of the device. Location-aware technology is often enabled in a device through the integration of electronic modules, ranging from a full GPS device that can be installed in vehicles to small GPS microchips integrated in mobile devices such as phones and tablets. The location-aware technology in the device interfaces with the operating system of the device to deliver location-aware or location-specific services.

Location-aware network has been wildly used. Location information is becoming an integral part of different mobile devices. Current mobile services can be enhanced with location-aware features, thus providing the user with a smooth transition towards context-aware services. Potential application fields can be found in areas such as travel information, shopping, entertainment, event information and different mobile professions. Some applications for Android and iOS feature different data and content based on the location of the devices, such as showing different events or tourist destinations, landmarks and even restaurants. Large campuses can use location-aware technology to guide students, especially newcomers, to specific locations such as classrooms and laboratories. The technology is increasingly being used in location guidance, especially in vehicles (GPS navigation) and for directing tourists in new cities.

A mobile phone can be located by the telecom operator in the network. The positioning is based on identifying the mobile network cell in which the phone is located, or on measuring distances to overlapping cells. In urban areas the accuracy can be down to 50 meters. Considering the issue of security, if the user wants to use location-aware services from other service providers, the location has to be transferred to the other service provider and the telecom operator must to get permission for this from the user [77].

The hospital workers are highly mobile. Their work includes visiting patients, locating resources, such as medical records, or consulting with other specialists. The information required by these specialists is highly dependent on their location. Access to a patient's laboratory results might be more relevant when the physician is near the patient's location and not elsewhere [11]. Many chronic patients are highly mobile as well. They are constantly changing location to perform their daily life. It is important for the nearest hospital professions to access their EMR immediately once an emergency has occurred [78].

In the proposed system, we describe a location-aware medical data system that was developed to provide access to patients' EMR, based on patients' location (home area, work place, family's home, friends' home, and other places). Due to the various staying time of patient in different location and the limited storage capacity of femtocaching, the femtocaching would choose the most suitable medical files to cache. Once an emergency has occurred, the physician can access the patient's EMR with low delay. The purpose is to provide a comprehensive medical service for patients no matter where the patient is.

3.1.3 Overview of Medical Video and Dynamic Vision Sensor

Medical video is a significant part of the health care area. Video was selected to provide important reference to physicians and an additional option for close supervision and offered increased privacy and as a result, decreased stress for the patients [79]. In [80], video has been used to monitor the patients with sleep disorder or sleepwalking. The sleeping behaviour of patients is recorded by camera at night. According to the monitoring video, physicians could provide effective medical services to patients. These results suggest that video captures most of the important features of the disorder. Video monitoring is proposed to reduce the number of falls for patients as well. Through remote monitoring of patients via video surveillance technology, trained hospital staff can observe the activity of multiple high-risk patients from a central location on each floor. The results show that video monitoring could significantly reduce the number of falls and more importantly, the injuries they often cause. For patient privacy, the video is live and not recorded. Used in conjunction with advanced communications and other procedures, video monitoring can lower staffing requirements while reducing risk and exposure for patients and the facility itself [81]. However, the size of medical video is huge and has significant data redundancy issues. For example, monitoring the patient who has sleep disorder or sleep walking condition, the effective part of monitoring is when patient starts to move or walk. However, the monitoring camera would keep recording during the whole night no matter patient is sleeping or walking. In this chapter, DVS technology is proposed to reduce the size of visual medical files. The transmission delay can be decreased when transferring the medical files among the hospitals. The clinical professionals could access the EMR within a shorter time and provide more efficient health care services.

DVS camera is different to the conventional video camera; it can detect temporal contrast of brightness and contains an array of asynchronous autonomous self-signaling pixels which respond quickly to relative changes according to the intensity of light [82]. The conventional video cameras are frame based, which means the camera captures a certain number of still images in one second and rapidly plays it back, giving the viewer the illusion of motion [20]. Conventional video cameras have two main disadvantages:

- The first being the limited number of frames which will miss most of movement if the objects move too fast [21]. For example using a conventional video camera with 30 frames per second to record the punches of a boxer; the camera cannot avoid the problem that all the image sensors have the same timing source. This weakness leads to inadequate data when analysing nuances of the boxer's arm motion. Higher frames and sophisticated post processing may improve this case. However, limited bandwidth, power and computing resources again limit this possibility.
- The second disadvantage is the huge storage capacity required due to the large quantity of data produced by these cameras. The real-time object tracking system based on video data processing requires large computational effort and is consequently done on high-performance computer platforms [22]. For example in using a conventional video camera to monitor and record the sleep behavior pattern of a narcolepsy patient at home, only the data relating to the patients' movements are needed. However, huge unvalued and redundant data where the subjects are not moving including the furniture and other unchanging parts of the room is recorded and processed by the monitoring system, which significantly increases processing time and causes a waste of storage space and transmission bandwidth.

Compared with the conventional camera, the DVS camera has the fastest response time in recognising fast motion. As mentioned before, the conventional camera is frame-based. However, the DVS camera is event-based. Sampling with different rates along with the intensity changing of light not only provides more accurate analysis for the fast motions, but also decreases the use of storage capacity, bandwidth and power [20].

A DVS camera typically requires orders of magnitude lower storage capacity than that of a conventional frame-based camera [72]. For the working pattern of DVS, when subjects move faster, DVS camera takes more pictures per second.

When subjects move slower, DVS camera takes less pictures per second. If there is no change of subjects, DVS camera does not take any pictures.

In [19], the authors use DVS camera to track the real-time movement of vehicles. The sensor outputs a measurement only when the light striking a pixel shifts by a present amount. When the light level crosses that threshold, circuitry attached to the pixel produces a “change-detection event” and triggers an illuminance measurement, encoded by the interval between two more events. The results show it can reduce the computation burden significantly compared to traditional traffic surveillance systems. In [73], it shows a demonstration of the algorithm on simulated DVS camera for a person tracking application. The data has been simulated from a video sequence with a resolution of 140×180 pixels that is close to developed 128×128 imager. Authors presented an embedded vision system and tracking algorithm using data from an asynchronous transient vision sensor for real-time applications. The results show that the tracking algorithm benefits from the capability of this imager to detect relative intensity changes and on its efficient asynchronous communication, which significantly reduces the computational burden as compared to traditional video-based traffic surveillance systems, enabling a low cost, low power consuming, and high bandwidth utilization implementation.

3.1.4 Three Levels of Electronic Medical Records

In a standard EMR, the patients’ data regarding all diagnosis, symptoms, prescribed medications, previous tests, and future anticipations are stored [83]. In our proposed system, EMR is defined into three levels according to the type and size of the files. The first level is a text/word type document, which contains all the basic information of patients, medical history, and conversation between physician and patient, etc. The second level corresponds to digital image files, which gives an idea of the patient’s disease status. The digital images include X-rays, ultrasounds, and

MRI (Magnetic Resonance Imaging), etc., which provide an immediate and directive information to physician. The third level is video and audio type, which mostly records the daily behavior or living habit of patient, especially for the patient monitoring.

Table 3.1: Three Levels Electronic Medical Records

Three Levels EMR	The Type of Document
The First Level	Text/Word
The Second Level	Digital Image
The Third Level	Video and Audio

- First Level: identification sheet, problem list, progress record, future anticipation, hospital administration records, email between medical professions and patients, billing records. The size of each file is up to 500MB.
- Second Level: X-rays (up to 2GB), ultrasounds (up to 5GB), Magnetic Resonance Imaging (MRI, up to 10GB), Computed Tomography (CT, up to 15GB), Positron Emission Tomography (PET, up to 20GB), Computed Radiography (CR, up to 15GB).
- Third Level: Medical Video records (up to 100GB), Medical Audio records (up to 100 GB).

As we can find, the size of the second and third level is much larger than text and word files. At the moment, most hospitals use a separate storage device, such as a mobile hard disk, to cache medical images. However, in terms of patients, they could lose them easily. For the physician, it becomes really difficult to access the medical images again to provide effective medical services. For the medical video/audio, the transmission delay is a big challenge, especially monitoring patient online. In this chapter, we proposed a new system to allocate and cache the three levels EMR and decrease the transmission delay when physicians access to medical files.

3.1.5 Contributions and Structure of Chapter

In the previous study [17] [18], femtocaching is proposed to receive and cache popular videos of new media platform. For the users who want to watch the videos, they can download from femtocaching instead of macro base station, thus decreasing the burden of networks. In our research, we allocate patient's medical records to the nearest hospitals of home area, work place, family home, friends' home, and other places according to the patient's social life. Location-aware femtocaching has been used to caching patients EMR and set up close to all the hospitals. The recently developed dynamic vision sensor technology is exploited to achieve smaller medical visual data files in our proposed tele-health system. DVS camera could improve the bandwidth utilization and save more resources, such as devices battery life and storage capacity of camera to current monitoring system, especially for the patients monitoring. The DVS camera only takes video when patient is moving. It is the first time that femtocaching technology and DVS camera are proposed in healthcare area.

We divided electronic medical records into three levels according to the type and the size of medical files (word/text, image, and video/audio). The EMR is allocated in femtocaches efficiently and effectively according to our proposed medical records allocation algorithm by applying knapsack model with penalty minimization. Due to the limited storage capacity of caches, not all the caches have the enough space for medical files, especially medical images and videos. The proposed medical records allocation algorithm would choose the most suitable files to cache based on three important penalty parameters including the staying time of patient in one location, the importance of medical records to patient's current health condition, and the transmission delay of medical files. The purpose is to provide the medical files in a timely manner and reduce delay when medical files are being shared and allocated among the hospitals. The simulation results show that the proposed scheme can drop the transmission delay of EMR by 43.57% to 86.66% compared with the scenario without using location-aware femtocaching technology. We also compared the simulation results of transmission delay with DVS camera technology, which could decrease the transmission delay by 41% to 80.4% more compared to the scenario with femtocaching only, which means the transmission delay can be improved by 88.94% to 92.13% comparing with the scenario without proposed system. It is a significant improvement comparing to traditional EMR transmission.

The rest of the chapter is organized as follows. In Section 3.2, we provide the system model with location-aware femtocaching based on patient's daily location. Section 3.3 focuses on the medical records allocation algorithm and system analysis. This is followed by the three penalty parameters and transmission delay calculation. In Section 3.4, a comprehensive performance evaluation has been conducted. Finally, Section 3.5 concludes the chapter.

3.2 System Scenario and Modeling

A new system is proposed to allocate patient EMR by using femtocaching and knowledge of patients' location. In the proposed scenario we assume that patients have wearable mobile devices with positioning facility to record their locations, and the registered hospitals have the complete EMR of the respective patients.

3.2.1 Location-aware Wearable Sensors for Health Condition Tracking

A new system is proposed to decrease the medical video size, transfer, allocate and cache patient medical records by using DVS camera and location-aware femtocaching technologies. In the proposed system, we assume that patients use location-aware wearable sensor devices to monitor their health condition and record their location at real time. The wearable sensor has two main functions in our scenario.

- The first one is the detection of emergency situations (e.g. falls). In an emergency situation, an electronic impulse produced by patient's body may be above or below the critical value. The sensor would then send an emergency message to the hospital and clinical professionals. The patients can receive immediate medical services [84]. The caregivers and family of the patient are notified when the emergency situation happens. Other notifications would be sent in many cases, for example, if the patient requires assistance in taking their medicine. Clinical personnel can remotely monitor patient's status and be alerted in case a medical decision has to be made.
- The second main function of the sensor is the recording of the patient's location. Using this approach, a warning about the fall and the location of the subject undergoing monitoring is transmitted to a caregiver or family member via SMS, email and Twitter messages, etc. This approach can record how long the patient stays in this location, which provides an effective reference to allocate medical records to the femtocaching of the nearest hospital.

3.2.2 The Proposed New System by Applying Femtocaching and DVS Camera

Firstly, we consider a scenario with one location, as shown in Figure 3.2, Femtocaching A and Femtocaching R are located in the nearest hospital of patient's home area and the registered hospital respectively. We assume that the registered hospital has the complete EMR of the patients. Femtocaching A downloads medical records from registered hospital during low data traffic times, e.g. at night, then allocates and caches the suitable medical files by using the proposed medical records allocation algorithm.

Considering the limited storage capacity, the Femtocaching A would cache the most important suitable medical records by applying the allocation algorithm. Due to the close distance between Femtocaching A to the home hospital, the respective physician can access patient's medical files very efficiently. In this monitoring system, we install DVS camera and wearable sensors in the home area (e.g. for tracking the sleeping pattern of a patient suffering from sleep disorder). The data produced by DVS cameras and wearable sensors are transmitted to Femtocaching A. The nearest hospital of home area would update all the new medical records to registered hospital during the low data traffic times as well. The system would ensure the registered hospital keep the latest medical records for patients.

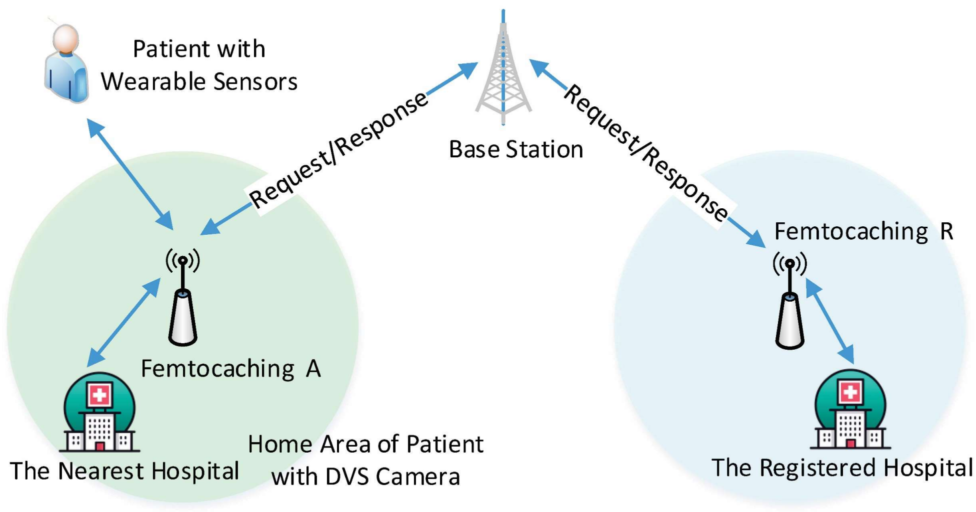

Figure 3.2: The Scenario of Femtocaching Applied in Home Area.

Here we consider a typical case where a patient each day on average spends 10 hours in the home environment, 8 hours in workplace, 3 hours for visiting relatives, 2 hours for visiting friends and 1 hour at other locations. The scenario is shown in Figure 3.3. The probabilities for spending time at each location are 41.67%, 33.33%, 12.5%, 8.33% and 4.17%, respectively. Same as the home environment, the DVS camera and wearable sensors are located in workplace of patient (e.g. for tracking the diet information and blood sugar level for a patient who is suffering diabetes). In the family's home, friend's home, and other places, patients use wearable sensors to record current health situations. Femtocaching *C*, Femtocaching *D*, and Femtocaching *E* are set up near the hospitals for visiting relatives, friends and other places respectively. All the hospitals in different location receive, allocate, and cache the medical records from the registered hospital. The real time medical data from wearable sensors and DVS camera would be sent to the nearest femtocaching. The physicians could provide effective medical care to the patients once the emergency situation occurs. The latest medical records would be updated to the registered hospital after patient receives health care services from one of the hospitals. The system would ensure the registered hospital keeps the latest medical records for patients. Therefore, wherever the patient stays in, the nearest hospital can access the medical records and provide effective and efficient medical services to the patients by applying the proposed system.

In the proposed system, it is an ideal situation that physicians can provide effective medical services with the cached files in femtocaching. However, if physicians need more medical records which are not cached in femtocaching, the registered hospital has to send it to the nearest hospital through macro base stations, which costs more vital and precious time for patients and can decrease the health care quality. Therefore, having an effective medical records allocation algorithm is very important. The new allocation algorithm based on knapsack model with penalty minimization is proposed in next section to determine which medical files has higher priority to be stored in the femtocaching firstly.

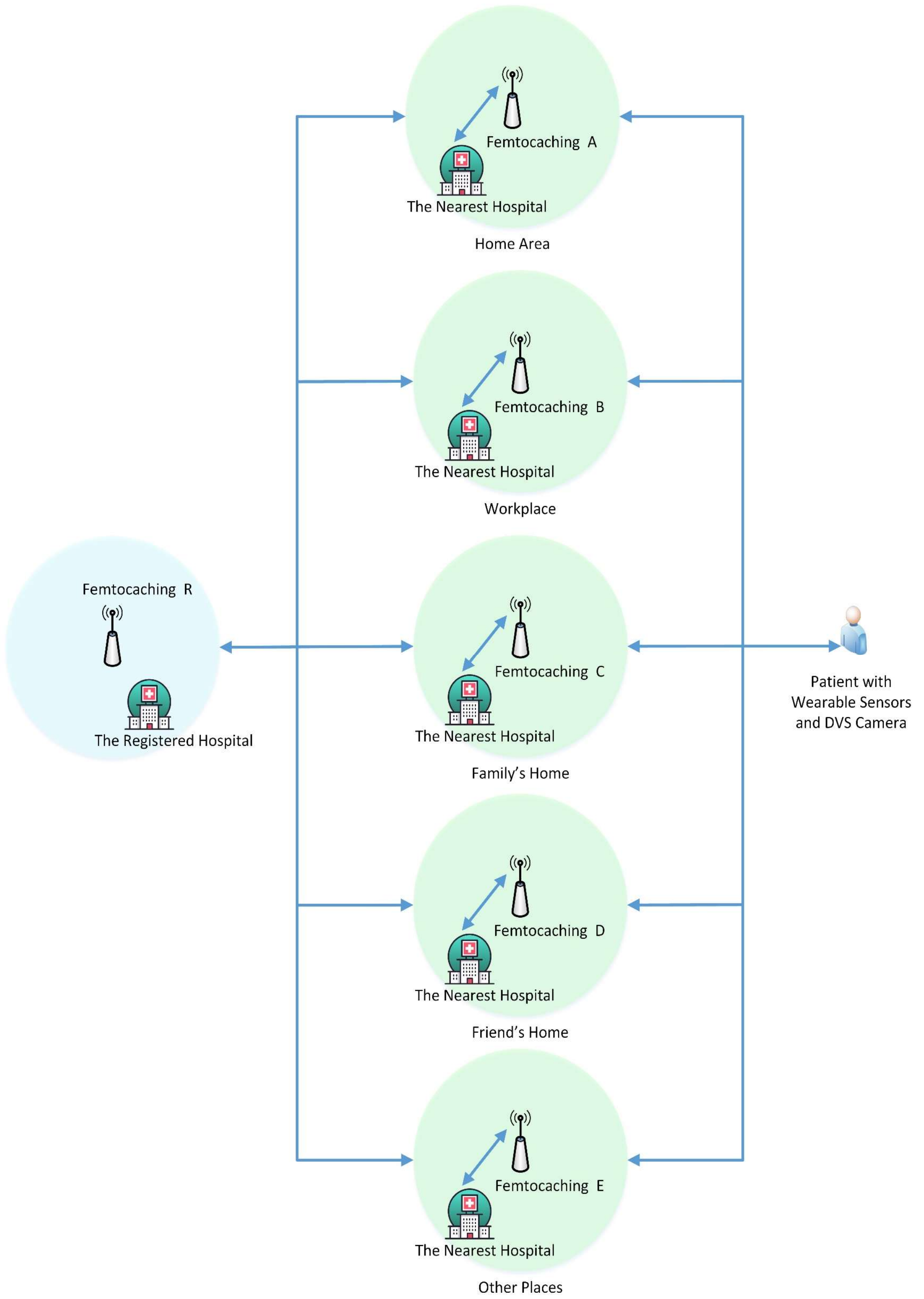

Figure 3.3: The Proposed Scenario Based on Patients' Daily Activities.

3.2.3 DVS camera for visual medical data

In our proposed system, video records account for most of the space of EMR. We study a scenario with limited edge caching capacity. We assume that video records require storage capacity of 200GB. However, the size of video files, such as those for recording elders' fall or monitoring alzheimer patients, can be much greater than 200GB, especially if a high-resolution camera is deployed. In the proposed system, we have used common international format（CIF）resolution DVS cameras as CIF resolution (352x240) is typically used by mid-level stand-alone digital video recorders (DVRs) when recording real time videos. It is also typically used by high-end systems for remote internet viewing specifically monitoring and surveillance cameras.

IniLabs [85], a company specialized in DVS technology, studied sleep behaviour pattern by using 128x128 DVS cameras. The DVS only outputs the subjects' movements. A whole night of sleep can be recorded in 100 MB of storage and played back in less than a minute. Activity levels can be automatically extracted, and any parts of the recording can be viewed at one millisecond resolution. The bit rate for 128x128 DVS cameras are suggested to be 256kbps.The bit rate of a CIF camera is 512kbps. Hence if we use 352x240 CIF with DVS to record the sleep behaviour pattern for a whole night, the size of the video will be around 200MB. However, the size of recorded visual data using conventional CIF resolution cameras is around 2.35GB (512x12x3600/8/1024) for a whole night.

3.3 Analytical Model with Femtocaching and DVS

In the first part of this section, the medical records allocation algorithm is proposed to choose the most suitable files to cache in femtocaching based on three important penalty parameters including the staying time of patient in one location, the importance of medical files to patient's current health condition, and the transmission delay of medical files. The optimised allocation results have been calculated. For the second part, we analysed the system. The medical record transmission delay is addressed.

3.3.1 Medical Records Allocation Algorithm

Due to the storage capacity limitation of the femtocaching, a medical records allocation algorithm is proposed to select the most suitable medical records stored in the femtocaching. We assume the capacities of femtocachings are different. In our case study, the capacities of Femtocaching A, B, C, D, and E are 100GB, 500GB, 100GB, 50GB, and 10GB, respectively. For the Femtocaching R, we assume it has enough storage to keep the whole EMR.

The wireless distributed caching network with femtocaching technology proposed by [8] selects the files according to popularity, which is not applicable to the healthcare environment setting due to ethical concerns. Therefore, we deploy knapsack model with penalty minimization [9] for optimizing the process of femtocaching file placement. Each part of EMR is correlated to a series of penalty parameters according to the patient's staying time in one location, the importance of the medical records in terms of the disease of the patient, and the size of the EMR files. The total value of all three penalty parameters would decide the priority of each type of files. In the proposed knapsack model, the optimization problem is to minimize the overall penalty in order to determine the EMR with the lowest penalty and highest priority to be stored in the femtocaching.

a). The staying time of the patient in one location

The staying time of the patient is one of the important penalty factors. The patients who stay longer in one area should be given higher priorities to cache medical files which are related to the disease of patient [86]. Therefore, the events of higher priority should be associated with lower penalties. The optimisation problem will be modelled by Equation (3.1), where α_f is equivalent to the staying time penalty parameters of patient in one area [87]:

Minimize:

$$P = \sum_{f \in M} \alpha_f \, x_f$$

Subject to:

$$x_f = 0 \text{ or } 1 \tag{3.1}$$

where

P: The total penalty.

M: The set of medical records including text/word, image, video/audio, and video/audio with DVS camera.

α_f: The penalty coefficient of choosing File f.

x_f: The decision variable of caching File f in femtocaching or not.

The wearable location-aware device would record the amount of time that a patient should stay in one location. Patient who has the smallest penalty parameters has the highest priority to cache medical records in home area. We consider that the sum of staying time by hours and penalty coefficient values is 24. Therefore, in the proposed system, the penalty coefficient of home area, workplace, family's home, friend's home, and other locations is 14, 16, 21, 22 and 23 respectively. The main purpose of this factor is to decide which femtocaching the system should cache at first place.

b). The importance of medical records

According to the disease condition, the importance of medical records should be considered for placement algorithm. For example, the patients with fractures, the X-ray images are more important than video and audio files in terms of diagnosing the disease [88]. Therefore, the EMR allocation algorithm should associate higher priority to the images file. The updated optimization model in Equation (3.1) after inclusion of the penalty parameters λ_f corresponding to the importance of medical records is shown below:

Minimize:

$$P = \sum_{f \in M} \alpha_f \, x_f + \sum_{f \in M} \lambda_f \, y_f$$

Subject to:

$$x_f, y_f = 0 \text{ or } 1 \tag{3.2}$$

where:

λ_f: The penalty parameters of the importance of medical records.

y_f: The decision variable of caching File f in femtocaching or not.

We assume that image file is the most important part according to the patient's disease condition and the video/audio file is the least important part in our proposed system. Therefore, the medical image files have the highest priority to cache and have the smallest penalty coefficient. The penalty coefficient λ_f of three type of medical files can be presented as in Table 3.2. The main purpose of this factor is to decide which medical file that femtocaching should cache at first place.

Table 3.2. The Importance of Medical Files Penalty Coefficient.

File Type	Penalty Coefficient λ_f
Images	1
Text/word	2
Video/Audio, Video/Audio with DVS	3

c). Transmission delay of medical records

Transmission delay of medical records is one of the important quality factors in proposed system. The ideal solution for minimizing access delay to patients' records is to retain the entire EMR in all local femtocachings. However, considering the capacity storage limitations and the increasing medical records of patient, the full EMR cannot be stored in all local caches. Therefore, an EMR allocation algorithm is required to achieve the minimum access delay, given the file priority and storage capacity constraints. The more medical records are cached in femtocaching, the less possibility that physicians require more medical files from the registered hospital, the higher priority to cache more medical files. The main purpose of this factor is to encourage patients to cache their EMR as much as possible in each femtocaching. The updated optimisation model in Equation (3.2) after the inclusion of the penalty parameters of the transmission delay of medical records (β_f) and limited storage constraint is shown below:

Minimize:

$$P_n = \sum_{f \in M} \alpha_f \, x_f + \sum_{f \in M} \lambda_f \, y_f + \sum_{f \in M} \beta_f z_f$$

Subject to:

$$\sum_{f \in M} \varphi_{fn} y_{fn} < S_n$$

$$x_f, y_f, z_f = 0 \text{ or } 1 \tag{3.3}$$

where:

β_f: The penalty parameters of the transmission delay.

z_f: The decision variable of caching File f in femtocaching or not.

φ_{fn}: The size of the medical files that cache in Femtocaching n.

S_n: The storage limit of each femtocaching.

As mentioned in the previous section, the patient's EMR is defined into text/word, images, and video/audio. Therefore, there are 7 combinations including text/word, images, text/word and images, video/audio, video/audio and text/word, video/audio and images, the whole EMR (text/word, images and video/audio). Caching more EMR files in local femtocaching means less transmission time and a lower penalty. Higher priority is equivalent to lower penalty. The penalty coefficient β_f can be presented as in Table 3.3. Due to the decreased size of medical video by applying DVS camera, the order of the combinations is different than the previous one. The penalty coefficient by applying DVS camera is shown in Table 3.4.

According to the capacity of each femtocaching, and the proposed three penalty coefficients, the optimization model based on knapsack model with penalty minimization in Equation (3.3) is implemented by using LpSolve. The optimized results with femtocaching only and applying DVS camera are shown in Table 3.5 and Table 3.6 respectively.

Table 3.3. Penalty Parameters of Transmission Delay.

Caching Combinations	Penalty Coefficient β_f
Text/word, images, video/audio (290G)	1
Images, video/audio (287G)	2
Text/word, video/audio (203G)	3
Video/audio (200G)	4
Text/word, images (90G)	5
Images (87G)	6
Text/word (3G)	7

Table 3.4. Penalty Coefficient of Medical Records by Adding DVS Camera.

Allocation Combinations	Penalty Coefficient β_f
Text/word, images, video/audio with DVS (106.66G)	1
Images, video/audio with DVS (103.66G)	2
Text/word, images (90G)	3
Images (87G)	4
Text/word, video/audio with DVS (19.66G)	5
Video/audio with DVS (16.66G)	6
Text/word (3G)	7

Table 3.5: The Optimised Medical Files Allocation.

Femtocaching	Medical files allocation
Femtocaching *A*	Images, Text/words (90G)
Femtocaching *B*	Text/word, Images, Video/Audio (290G)
Femtocaching *C*	Images, Text/words (90G)
Femtocaching *D*	Text/words (3G)
Femtocaching *E*	Text/words (3G)

Table 3.6: The Optimised Medical Files Allocation with DVS.

Femtocaching	Medical files allocation
Femtocaching *A*	Images, Text/words (90G)
Femtocaching *B*	Text/word, Images, Video/Audio with DVS (106.6G)
Femtocaching *C*	Images, Text/words (90G)
Femtocaching *D*	Text/words, Video/Audio with DVS (19.6G)
Femtocaching *E*	Text/words (3G)

3.3.2 Simulation of Transmission Delay

The medical record transmission delay [89] in the proposed system can be presented as follows when patient is in home area:

$$D_T = \frac{N_1}{R_1} + \frac{N_2}{R_2} \tag{3.4}$$

where

D_T: The transmission delay.

N_1: The medical records which are cached in femtocaching in bits.

N_2: The rest of medical records in bits.

R_1: The rate of transmission of femtocaching (bits per second).

R_2: The rate of transmission of macro cellular network (bits per seconds).

Therefore, the average transmission delay of medical records with home area, workplace, visit relative, visit friends, and other places in the proposed system can be presented as (3.5):

$$D = P_H D_{TH} + P_W D_{TW} + P_{FM} D_{TFM} + P_{FD} D_{TFD} + P_O D_{TO}$$

$$= P_H \left(\frac{N_{H1}}{R_1} + \frac{N_{H2}}{R_2}\right) + P_W \left(\frac{N_{W1}}{R_1} + \frac{N_{W2}}{R_2}\right) + P_{FM} \left(\frac{N_{FM1}}{R_1} + \frac{N_{FM2}}{R_2}\right) + +P_{FD}(\frac{N_{FD1}}{R_1} + \frac{N_{FD2}}{R_2}) + P_O(\frac{N_{O1}}{R_1} + \frac{N_{O2}}{R_2}) \tag{3.5}$$

where

P_H , P_W , P_{FM} , P_{FD} , P_O : The possibility that patients stay in their home area, workplace, visit relative, visit friends, and other places respectively.

D_{TH} , D_{TW} , D_{TFM} , D_{TFD} ,D_O : The transmission delay of home area, workplace, visiting family, visiting friends and other places.

N_{H1}, N_{W1}, N_{FM1}, N_{FD1}, N_{O1}: The size of medical records which are cached in Femtocaching *A*, *B*, *C*, *D*, and *E* in bits.

N_{H2}, N_{W2}, N_{FM2}, N_{FD2}, N_{O2}: The size of the rest of medical records in bits.

We used Poisson Distribution [90] to simulate the possibility of patient's routine activities. We calculated the delay for the case where number of occurrences increases to 1000. Therefore, the average transmission delay of medical records can be presented as:

$$D = \frac{\sum_K \frac{e^{-\lambda_H} {\lambda_H}^{K_1}}{K_1!}}{K} \left(\frac{N_{H1}}{R_1} + \frac{N_{H2}}{R_2}\right) + \frac{\sum_K \frac{e^{-\lambda_W} {\lambda_W}^{K_2}}{K_2!}}{K} \left(\frac{N_{W1}}{R_1} + \frac{N_{W2}}{R_2}\right) + \frac{\sum_K \frac{e^{-\lambda_{FM}} {\lambda_{FM}}^{K_3}}{K_3!}}{K} (\frac{N_{FM1}}{R_1} + \frac{N_{FM2}}{R_2}) + \frac{\sum_K \frac{e^{-\lambda_{FD}} {\lambda_{FD}}^{K_4}}{K_4!}}{K} (\frac{N_{FD1}}{R_1} + \frac{N_{FD2}}{R_2}) + \frac{\sum_K \frac{e^{-\lambda_O} {\lambda_O}^{K_5}}{K_5!}}{K} (\frac{N_{O1}}{R_1} + \frac{N_{O2}}{R_2})$$

Subject to:

$$K_1, K_2, K_3, K_4, K_5 \in \{0, 1, 2, \ldots\ldots, 22, 23, 24\}$$

$$K_1 + K_2 + K_3 + K_4 + K_5 = 24 \tag{3.6}$$

where

λ_H: The average possibility that patients stay in their home area.

K_1, K_2, K_3, K_4, K_5: The staying time of patient in home are, workplace, family's home, friend's home, and other places respectively.

K: The number of occurrences.

λ_W: The average possibility that patients stay in workplace.

λ_{FM}: The average possibility that patients visit family.

λ_{FD}: The average possibility that patients visit friends.

λ_O: The average possibility that patients visit other places.

The other simulation parameters and conditions of the system are shown as follows. The simulation results are analysed by MATLAB and presented in the following section.

- The EMR allocation scenario is covered by a single 3GPP LTE R8 cell. The femtocaching uses a simplified 802.11n protocol.
- The femtocaching is set up near hospital within 100 meters.
- The values of λ_H, λ_W, λ_{FM}, λ_{FD}, and λ_O are 0.4167, 0.3333, 0.125, 0.0833, and 0.0417 respectively.
- The value of N_{H1}, N_{W1}, N_{FM1}, N_{FD1}, and N_{O1} is shown in Table 3.5 and Table 3.6. For the scenario of femtocaching only, the value is 90G, 290G, 90G, 3G, and 3G respectively. After adding DVS, the value is 90G, 106.6G, 90G, 19.6G, and 3G respectively.
- The value of N_{H2}, N_{W2}, N_{FM2}, N_{FD2}, and N_{O2} is the rest of EMR. For the scenario of femtocaching only, the value is 200G, 0G, 200G, 287G, and 287G

respectively. After adding DVS, the value is 16.6G, 0G, 16.6G, 87G, and 103.6G respectively.

3.4 Performance Evaluation

The results are studied in two different situations. The first, namely the best situation, corresponds to the case where the physicians have enough medical files from femtocaching to provide effective medical records and no need to access the rest of EMR at the registered hospital. In the proposed system, poisson distribution is implemented. We compared with the three simulation results including the scenario with femtocaching and DVS camera, femtocaching only, and without applying proposed technologies. The transmission delay of each femtocaching is shown in Figure 3.4. The transmission delay of Femtocaching *A*, *B*, *C*, *D*, and *E* without applying proposed system is 76.8, 247.47, 76.8, 2.56, and 2.56 minutes respectively. However, the transmission delay with femtocaching only is 10.24, 33, 10.24, 0.34, and 0.34 minutes respectively. The improvement of transmission delay is distinct in each femtocaching. The transmission delay is even improved more by adding DVS camera, which is 10.24, 12.13, 10.24, 2.34, and 0.34 minutes respectively.

The average transmission delay for the best situation is shown in Figure 3.5. The average transmission delay without proposed system is 124.4 minutes. However, the average transmission delay of femtocaching only is 16.59 minutes. For the scenario of adding DVS, the delay is 9.79 minutes. The proposed system with femtocaching only can drop the EMR transmission delay by 86.66%. However, after applying DVS, transmission delay can be decreased 41% more comparing with the scenario with femtocaching only. Therefore, the proposed system with femtocaching and DVS the can drop the EMR transmission delay by 92.13% in the best situation.

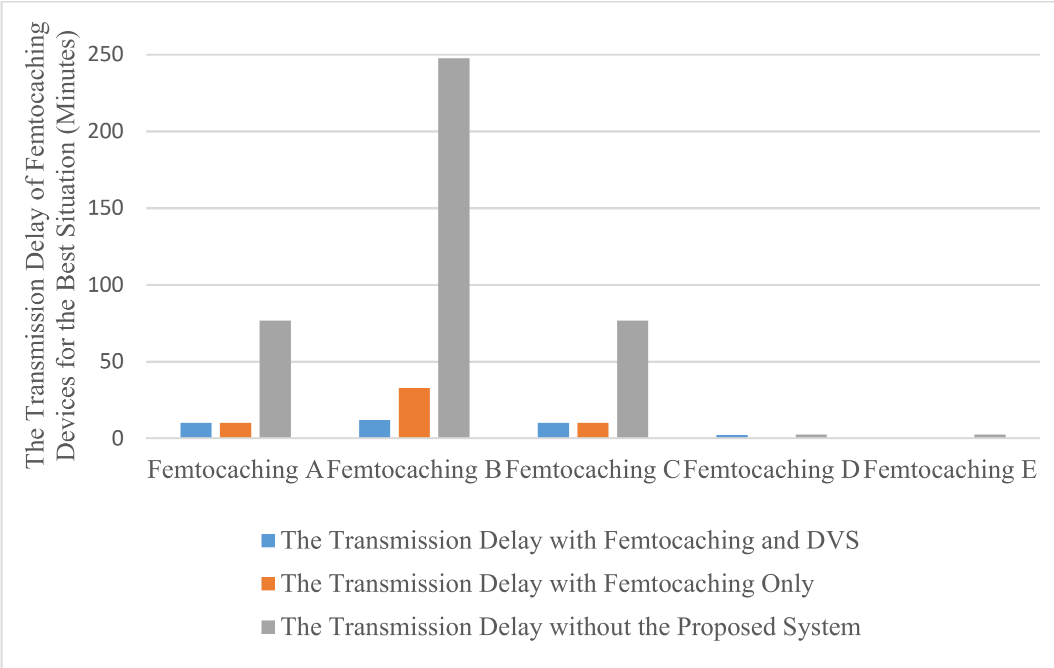

Figure 3.4. The Transmission Delay of Each Femtocaching for the Best Situation When the Number of Occurrences is 1000.

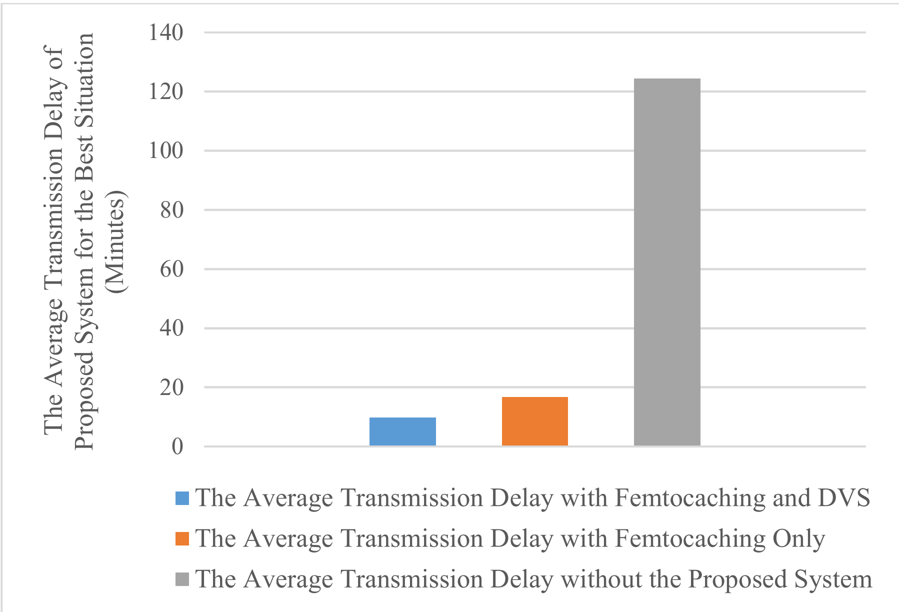

Figure 3.5: The Average Transmission Delay for the Best Situation.

The second one is the worst situation, corresponding to the case where the medical files cached in femtocaching cannot meet the requirement, and the physicians need to access the rest of EMR from the registered hospital. The transmission delay would increase when asking for the medical files through macro base stations. The transmission of each femtocaching is shown in Figure 3.6. The transmission delay of each femtocaching without proposed system is 247.47 minute. However, the transmission delay with femtocaching only is 180.9, 33, 180.9, 245.25, and 240.25 minutes respectively. The transmission delay is improved significantly by applying DVS, which is 24.46, 12.13, 24.46, 76.48, and 88.8 minutes respectively.

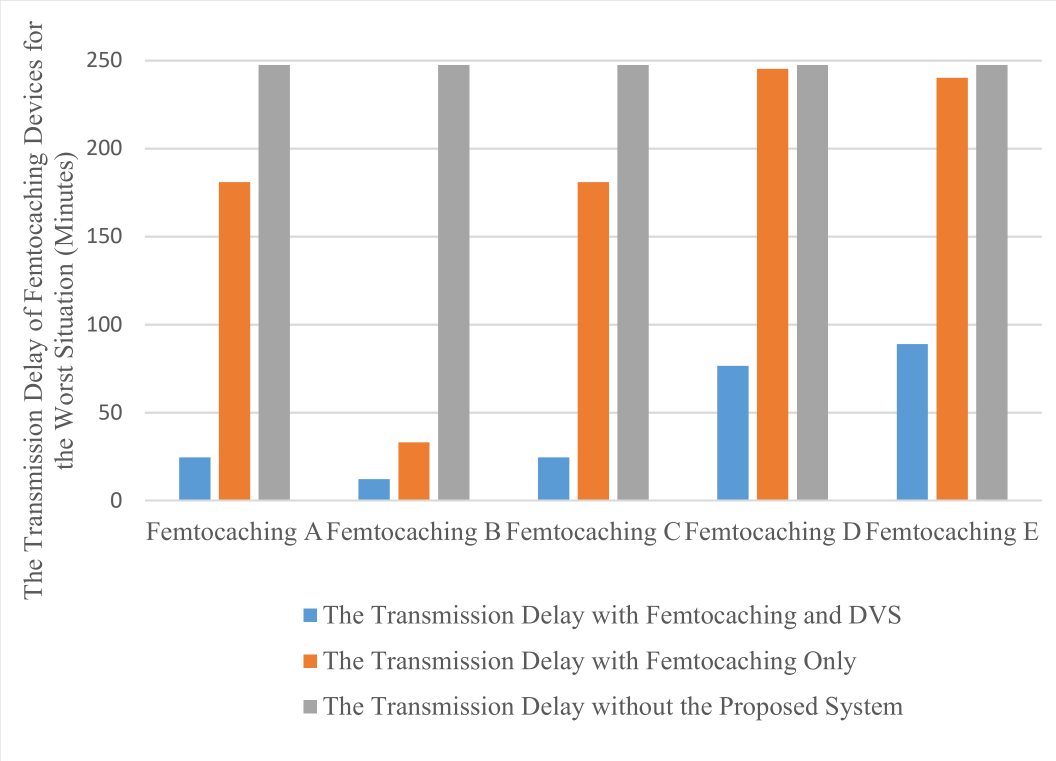

Figure 3.6: The Transmission Delay of Each Femtocaching for the Worst Situation When the Number of Occurrences is 1000.

The average transmission delay for the worst situation is shown in Figure 3.7. The average transmission delay without proposed system is 247.47 minutes. However, the average transmission delay of femtocaching only is 139.65 minutes. For the scenario of adding DVS, the delay is 27.37 minutes. The proposed system with femtocaching only can drop the EMR transmission delay by 43.57%. However, after applying DVS, transmission delay can be decreased 80.4% more comparing with the scenario with femtocaching only. Therefore, the proposed system with femtocaching and DVS the can drop the EMR transmission delay by 88.94% in the worst situation. Even though the transmission delay is not improved as much as in the best situation, the delay in is still less than the system without using our proposed scheme.

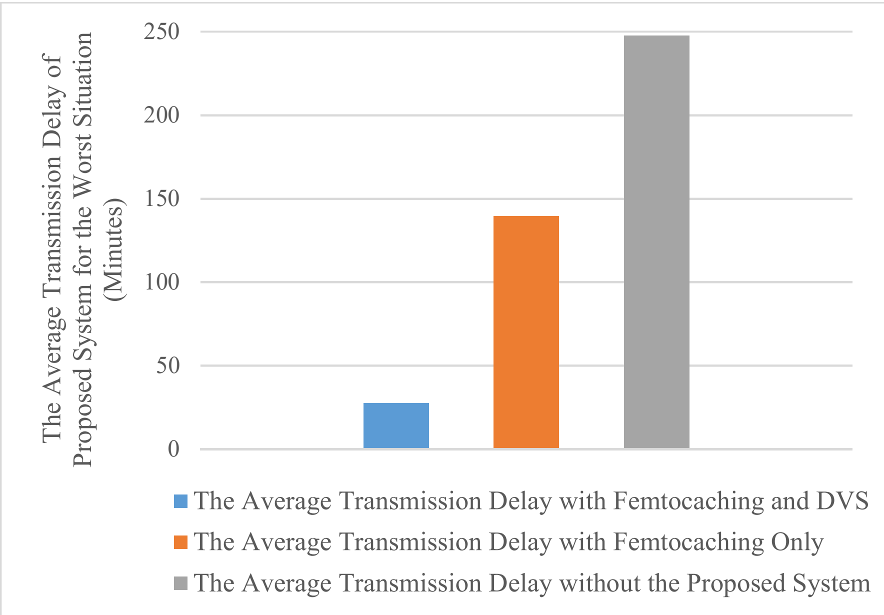

Figure 3.7: The Average Transmission Delay for the Worst Situation.

Therefore, the proposed scheme with femtocaching only can drop the average transmission delay of EMR by 43.57% to 86.66% compared with the scenario with traditional transmission method. The proposed scheme with femtocaching and DVS can decrease transmission delay by 88.94%-92.13%.

3.5 Summary and Conclusions

In this Chapter, location-aware femtocaching has been proposed to allocate and cache patient medical records. We present a system that provides timely medical services for patients based on their location (Home area, work place, family's home, friends' home, and other places). A new method is proposed to represent and structure patients' EMR according to the type and the size of medical records. We also propose an EMR allocation algorithm to allocate the most suitable medical files in femtocaching according to the patient's staying time in one location, the importance of the medical records in terms of the disease of the patient, and the size of the EMR files. The simulation results show that proposed scheme can drop the average transmission delay of EMR by 43.57% to 86.66% compared with the scenario without using location-aware femtocaching. In other words, the physicians could save up to 86.66% time to access to patient's medical records and provide effective health care services. In the proposed new system, a new medical video capture method is proposed to reduce the size of visual medical files by applying DVS camera technology. A new scenario has been proposed to transmit, allocate and cache patients' EMR (including visual data) by applying femtocaching. The transmission delay in the proposed system can be decreased significantly when transferring or sharing the medical files among the hospitals. The clinical professionals and patients can access the EMR within a short time. The simulation results show that the proposed scheme by adding DVS camera technology can decrease the transmission delay by 41% to 80.4% more comparing with our previous work (applying femtocaching only), which means the transmission delay can be improved by 88.94% to 92.13%.

Chapter 4

Auction-based and Non-cooperative Game Theory with Edge Computing for Sharing and Caching EMR among Patients

4.1 Introduction

By implementing EMR, patients' medical data can be tracked over an extended period of time by multiple healthcare providers [70]. On the other side, EMR is universal, meaning that instead of having different charts at different healthcare facilities [11], a patient will have one electronic chart that can be accessed from any healthcare facility by using EMR software [3].

In the previous chapter, we proposed a new method to represent and structure patient's EMR. In order to provide timely medical care, femtocaching technology has been used to cache and allocate EMR for patients. The simulation result shows that the proposed system can decrease the transmission delay of EMR by 43.57% to 86.66%. Dynamic vision sensor has been used to reduce the size of medical or monitoring video. The results show that transmission delay can be improved by 88.94% to 92.13%. We assumed that patient spends 10 hours at home, 8 hours in workplace, 3 hours for visiting relatives, 2 hours for visiting friends, and 1 hour at other locations. However, due to the changing location of patients, it is very difficult to allocate and cache medical records to the nearest hospital timely. Once an emergency occurs in the location where the medical files haven't been cached, the physician cannot access medical records at efficient time. Therefore, in this chapter, edge computing is proposed to process and cache EMR, which is an advanced telecommunication technology that users can process their data on the edge before sending to the public cloud regardless of the limitation of changing location. The edge devices can be a mobile phone, Wi-Fi, and home gateway. The data will be processed or partly processed on the edge device before being sent to the cloud.

Comparing with traditional cloud computing, edge computing could off load partly or whole data processing from cloud and provide more efficient and safer services. Patients can control their own medical data, made by wearable sensors, on the edge. We used edge computing to allocate patient's EMR on the edge instead of the cloud. By applying edge computing, not only the privacy of medical data can be protected, but also can lighten greatly network loads and decrease delay for cloud services. On the other hand, in the last chapter, patients only can use their own femtocaching to cache and allocate medical records. However, a new scenario has been proposed for patients to share their edge caching to others. A new medical records allocation algorithm has been applied by using action-based non-cooperative game theory.

We considered the scenario that patients share their storage capacity on the edge device to others and other patients could get effective and efficient medical care by caching their medical files to the adjacent edge devices. In the proposed system, four medical factors including the severity of illness, the size of the medical records that patient wants to cache, the queueing time that patients have waited, and the rewards that patients got by sharing edge caching with other patients, have been combined with telecommunication channel capacity for the patients to compete the allocation and caching priority. In the proposed non-cooperative game, Nash Equilibrium has been presented as a desirable outcome. We analysed the scenarios with 6, 60, 600, 6000 and 60,000 patients. The simulation results show that the proposed system can decrease the transmission delay by 56.87% to 93.69% in the best situation (the cached medical records are enough for medical support) comparing the system without sharing edge device, which means the proposed system saves more than half time for clinical professors to provide efficient medical services. In the worst scenario (the cached medical records are not enough for medical support and whole EMR is required), the simulation results show that the proposed system can decrease the transmission delay by 25.94% to 57.75%. The results show that the more shared edge devices, the higher channel capacity patients can get as well, which means the proposed system encourages more patients to share their storage capacity of edge devices to others.

4.1.1 Overview of Cloud Computing

With the rapid development of processing and storage technologies and the success of the internet, computing resources have become cheaper, more powerful and more ubiquitously available than ever before. This technological trend has enabled the realization of a new computing model called cloud computing [91]. Cloud computing has recently emerged as a new paradigm for hosting and delivering services over the internet [23]. Cloud computing is attractive to business owners as it eliminates the requirement for users to plan ahead for provisioning and allow enterprises to start small and increase resources only when there is a rise in service demand.

Cloud computing is one of the important technologies in the healthcare sector. In [23], cloud computing is proposed to collect patients' data. The purpose of the proposed design is to reduce the possibility of typing mistakes in the process of medical data collection and to provide always-on, real-time data collection as opposed to the manual handling of patient information. In [24], cloud computing is proposed to share patient medical records. The aim in [24] is to provide flexible access for patients and different medical professionals.

Cloud computing has been widely applied in various areas and has proven to be an effective method to collect and process all the data in the cloud. For the terminal users (i.e. mobile users, computer users etc.), different types of terminal devices can access the cloud easily and share data via the cloud [92]. By applying cloud computing in healthcare area, it's a great convenience for patents, family and medical care workers who want to access EMR immediately. For the service provider (i.e. medical care provider), the cloud is highly scalable. Larger scale services can be expanded easily if more service is demanded [91]. Cloud computing can be easily implemented as well.

The paradigm of cloud computing applied in healthcare is shown in Figure 4.1. The medical data from various sensors (i.e. heartbeat sensor, blood pressure sensor, fall detector sensor, and blood sugar sensor, etc.) is collected by medical care gateway (i.e. smart phone or other mobile devices). Medical care gateway transfers all the raw medical data to the cloud. Medical data is processed and analysed in the cloud. After medical care providers and families received requirements from the cloud, instructions, suggestions and configurations would be sent back to the patients. Cloud computing could connect medical service providers, patients and family together and provide effective health care services.

However, there's an increase in the amount of data correlating with the rapidly increasing number of IoT (internet of things) devices. The speed of data transportation is becoming the bottleneck for the cloud-based computing paradigm [25]. For example, wearable sensors and monitoring cameras produce huge amounts of image or video data every second and add a heavy traffic burden to the current networks. Cloud computing, especially, is not efficient enough for those health devices that require very short response times, thus it could cause intolerable network latency. On the other hand, the data which is produced by wearable sensors or other health devices, is usually private and confidential. Transferring and processing patient information through public cloud would pose a breach of patient confidentiality.

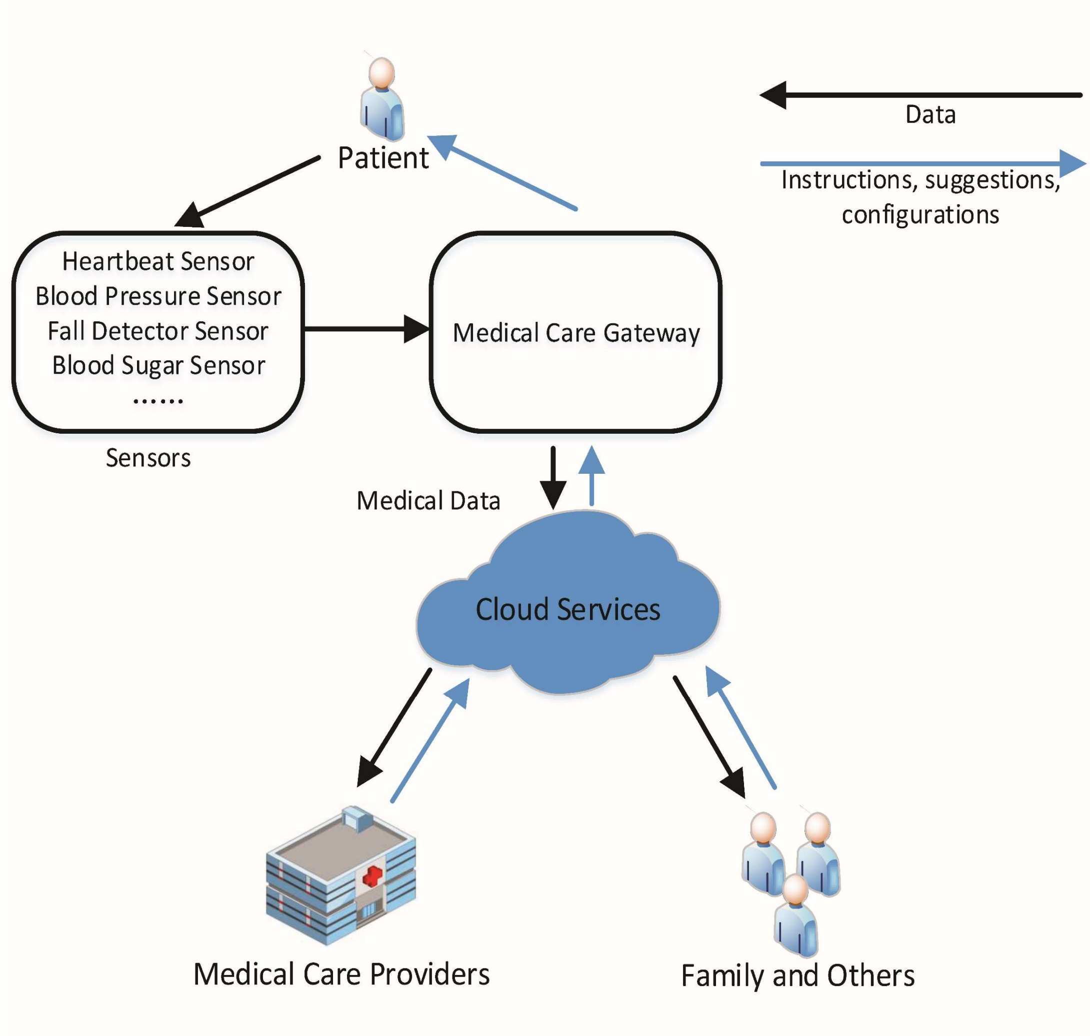

Figure 4.1: The Paradigm of Cloud Computing

4.1.2 Edge Computing in Health Care

Here we proposed edge computing to cache and share medical records. Edge computing enables mobile subscribers to access IT and cloud computing services at a close proximity within the range of radio access network (RAN) [53]. In the traditional cloud computing, all the information and data are transferred to the cloud network, then users would receive the data through the corresponding base station.

Comparing with the traditional cloud computing, edge computing has three main advantages:

- Edge computing could cache and process computing task at the edge of the network without transferring to the cloud network. Edge computing can offload part or all of network traffic from the cloud network to the edge, which would drop the network latency and decrease the bandwidth consumption [93].
- For the issue of privacy, caching patients' health data at the edge is safer than cloud. Patients can have their own medical information at a close proximity and control to decide if the health data should be transferred to cloud and service providers.
- Edge computing saves expense and does not need more network infrastructure. For a patient, mobile phone is the edge between body sensors and the cloud. For a smart home, gateway or Wi-Fi is the edge between home sensors and the cloud [25].

The paradigm of edge computing is shown in Figure 4.2. We consider the scenario where edge computing offloads all the network from cloud firstly. Raw medical data from various sensors is sent to edge devices. Before being sent to cloud, medical data is cached and processed at the edge. Medical care providers and family would send instructions, suggestions and configurations back after received requirements from cloud. In this scenario, cloud is used to transfer processed data only.

For the scenario that edge computing offloads part of the network from cloud, part of raw medical data is processed at the edge. The rest is sent to the cloud. Medical care providers and family send responses back after receiving requirements from the cloud. In this scenario, cloud is used to transfer processed data and process raw medical data both. Edge computing can increase the edge responsibility and allows computation and services to be hosted at the edge [54].

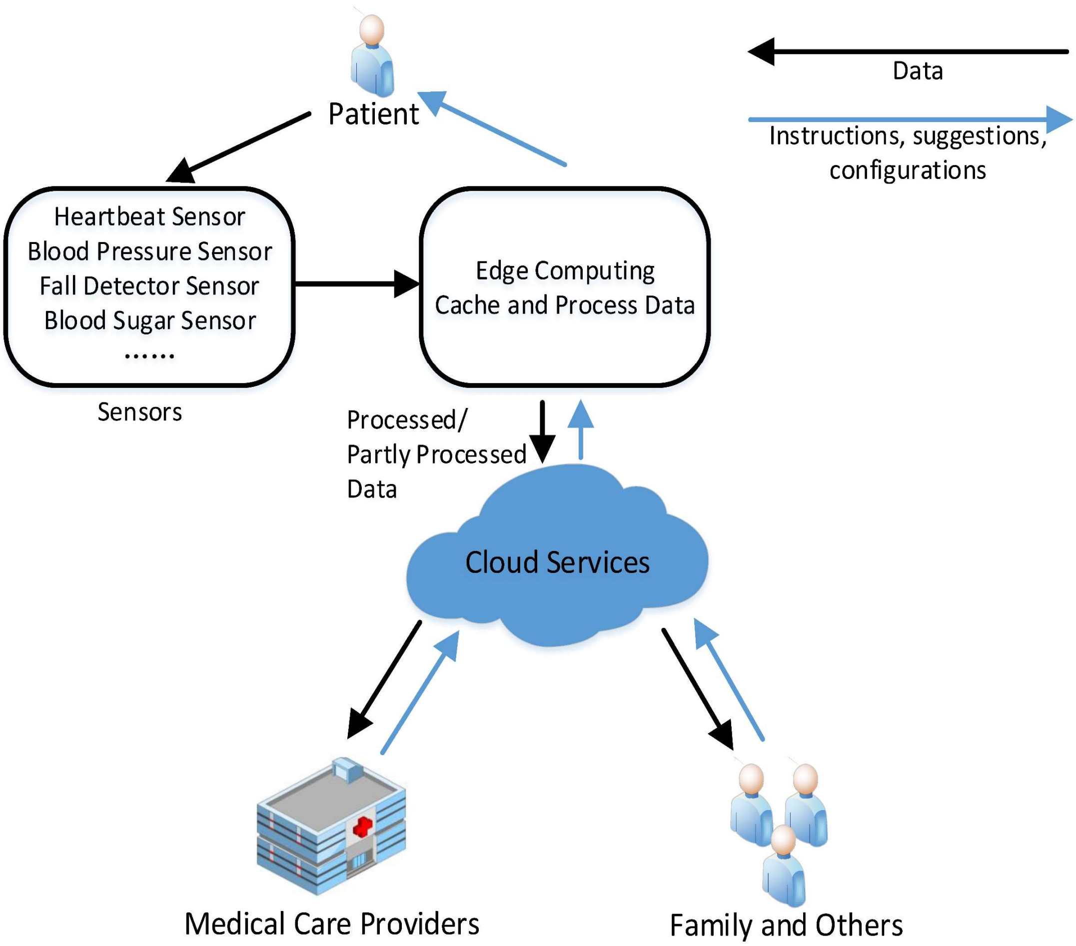

Figure 4.2: The Paradigm of Edge Computing

Until today, edge computing networks have not been deployed in cellular networks; thus, edge computing concept has been discussed only from a theoretical perspective so far. However, there are some related approaches that are similar to this concept. In [55], cloudlet-based offloading is proposed where mobile devices offload computational to the less resourceful server near the user proximity accessible using Wi-Fi access point. In [59], authors proposed FemtoClouds [25] system that provides a dynamic and self-configuring multiple devices cloud system to scale the computation of cloudlet by coordinating multiple mobile devices. In [94], authors proposed REPLISOM, an edge cloud architecture to reduce cloud responsiveness when multiple IoT devices replicate memory objects to the edge cloud through LTE environment.

4.1.3 Overview of Game Theory

Game theory provides useful mathematical tools to solve optimization problems with different competing players [11]. Game theory is a tool for analysing the interaction of decision makers with conflicting objectives. Economists have long used it as a tool for examining the actions of economic agents such as firms in a market [95]. Game theory has now become an important mathematical tool, which is used in situations that involves several entities whose decisions are influenced by the decisions of other entities playing with them.

Recently, game theory has been adopted in the telecommunication environment, especially in the wireless sensor networks, cognitive radio networks, and ad-hoc networks. Game theory is used as a tool for studying, modeling, and analyzing the interactions between individuals strategically [96]. In the wireless environment, game theory has been used in order to solve distributed power control, resource management and allocation, and dynamic pricing related problems [61]. It is concerned with finding the best actions for individual decision makers in these situations and recognizing stable outcomes.

Any game, when played, consists of the participants called players or agents of the game, each having his own preference or goal [97]. Each player of the game has an associated amount of benefit or gain which he receives at the end of the game, called payoff or utility, which measures the degree of satisfaction an individual player derives from the conflicting situation. For each player of the game, the choices available to them are called strategies. The solution of a game is referred to as Nash Equilibrium or Strategic Equilibrium, where each player cannot get a better payoff than the existing one by individually changing to another strategy. The utility function is a mapping of a player's choices into a real number. To understand the concepts presented so for, refer to Table 4.1, where two players *P1* and *P2* come in a strategic interaction to play this game. For ease, the game is represented by a matrix, called payoff matrix, which shows the choices available to players and the outcome for each choice he makes against the others. As shown in the payoff matrix of the game between *P1* and *P2*, each has to decide either to choose *X* or *Y*. The setup is strategic, so the choice of individual player's payoff depends on the choice made by other player as a response to his strategy. If player *P1* chooses *X*, then his payoff depends upon which choice the player *P2* makes. If *P1* chooses *X* and *P2* also chooses *X*, *P1* payoff will be 2, otherwise 10. Similarly, if *P1* chooses the *Y*, then depending upon the choice of *P2* he will either receive 10 or 3. The same choices and outcomes will happen for player *P2* as well. Here in this game, the players are *P1* and *P2*, the strategies for each are *X* and *Y* and the benefits or outcome of the game are represented in the form of matrix where each entry shows the payoffs in the form (payoff of *P1*, payoff of *P2*) duple. The numerical outcome for each player depends on the utility function used by each in the situation in which this game is played.

Table 4.1. Strategic Form of Game Theory.

P1 \ *P2*	*X*	*Y*
X	(2,8)	(10,5)
Y	(5,10)	(3,3)

The object of study in game theory is the game, defined to be any situation in which:

- There are at least two players. A player may be an individual, a company, a nation, a wireless node, or even a biological species.

- Each player has a number of possible strategies, courses of action he or she may choose to follow.
- The strategies chosen by each player determine the outcome of the game.
- Associated with each possible outcome of the game is a collection of numerical payoffs, one to each player.

According to the type of game, game theory can generally be categorized as non-cooperative and cooperative games.

a). Non-cooperative Game Theory

Non-cooperative game theory is concerned with the analysis of strategic choices and explicitly models the process of players' making choices out of their own interests [98]. Non-cooperative games can be classified into a few categories according to several criteria.

According to whether the players' moves are simultaneous or not, non-cooperative games can be divided into two categories: static and dynamic games. In a static game, players make their choices of strategies simultaneously, without the knowledge of what the other players are choosing. Static games are most often represented diagrammatically using a game table that is called the normal form or strategic form of the game. In the dynamic game, players involve strategic situations in which there is a strict order of play. Players take turns to make their moves, and they know what players who have gone before them have done. Dynamic games are most easily illustrated using game trees, which are generally referred to as the extensive form of a game. The trees illustrate all of the possible actions that can be taken by all of the players and also indicate all of the possible outcomes from the game.

According to whether the players have full information of all payoff-relevant characteristics about the opponents or not, the non-cooperative game can be classified into two types: complete information and incomplete information games. In the former each player has all the knowledge about others' characteristics, strategy spaces, payoff functions, and so on, but this is not so for the latter.

b). Cooperative Game Theory

A cooperative game (also called coalitional) is a game in which the players can make binding commitments, as is not the case in the non-cooperative game. Analysis in cooperative game theory is centered on coalition formation and distribution of wealth gained through cooperation. Within these two areas, finding procedures leading to outcomes that are most likely to occur under reasonable rationality assumptions in various game situations, and devising solution concepts showing attractive stability features are primary concerns in most research endeavors.

Cooperative game theory is most naturally applied to situations arising in political science or international relations, where concepts like power are most important. The definition draws the usual distinction between the two theories of games, but the real difference lies in the modeling approach. While in non-cooperative game theory the notion of the Nash Equilibrium is pervasive in capturing most aspects of stability, in cooperative game theory there is no solution concept dominating the field in such a way. Instead, there is a multiplicity of solutions, which is not due to the weakness of the theory, but rather to the inherent diversity of conflict situations into which it attempts to provide insight. Moreover, the main focus of the non-cooperative game is individual rationality and individual optimal strategy, but the cooperative game emphasizes collective rationality, fairness, effectiveness, etc., which means different things to different people [99].

The prisoner's dilemma is a standard example of a game analyzed in game theory that shows why two completely rational individuals might not cooperate, even if it appears that it is in their best interests to do so. The game is presenting as follows:

Two members of a criminal gang are arrested and imprisoned. Each prisoner is in solitary confinement with no means of communicating with the other. The prosecutors lack sufficient evidence to convict the pair on the principal charge, but they have enough to convict both on a lesser charge. Simultaneously, the prosecutors offer each prisoner a bargain. Each prisoner is given the opportunity either to betray the other by testifying that the other committed the crime, or to cooperate with the other by remaining silent. The offer is:

- If A and B each betray the other, each of them serves two years in prison.
- If A betrays B but B remains silent, A will be set free and B will serve three years in prison (and vice versa).
- If A and B both remain silent, both of them will only serve one year in prison (on the lesser charge).

Table 4.2: The Prisoner's Dilemma Based on Game Theory.

Prisoner *B* / Prisoner *A*	Prisoner *B* stays silent (cooperates)	Prisoner *B* betrays (defects)
Prisoner *A* stays silent (cooperates)	Each serves 1 year	Prisoner *A*: 3 years Prisoner *B*: goes free
Prisoner *A* betrays (defects)	Prisoner *A*: goes free Prisoner *B*: 3 years	Each serves 2 years

It is assumed that both prisoners understand the nature of the game, have no loyalty to each other, and will have no opportunity for retribution or reward outside the game. Regardless of what the other decides, each prisoner gets a higher reward by betraying the other ("defecting"). The reasoning involves an argument by dilemma: B will either cooperate or defect. If B cooperates, A should defect, because going free is better than serving 1 year. If *B* defects, *A* should also defect, because serving 2 years is better than serving 3 years. Hence either way, *A* should defect. Parallel reasoning will show that *B* should defect. Because defection always results in a better payoff than cooperation regardless of the other player's choice. Mutual defection is the only strong Nash Equilibrium in the game (i.e. the only outcome from which each player could only do worse by unilaterally changing strategy). The dilemma, then, is that mutual cooperation yields a better outcome than mutual defection but is not the rational outcome because the choice to cooperate, from a self-interested perspective, is irrational.

4.1.4 Overview of VCG Auction

The auction theory has recently been introduced to several types of resource allocation problems (e.g., time slot allocation, power allocation, and cooperative communication), especially spectrum sharing in telecommunication networks [100] [101] [102] [103]. Several types of auctions have been proposed and utilized in order to provide goods with users [101]. Among such auction mechanisms, it is well known that Vickrey-Clarke-Groves (VCG) auction mechanism satisfies incentive compatibility where the weakly dominant strategy for bidders is to bid truthfully [104]. In VCG auction mechanism, the auctioneer selects winners among all bidders so that social surplus can be maximized. VCG auction mechanism is extended to generalized VCG auction mechanism in order to consider multiple divisible units of goods.

A VCG auction is a type of sealed-bid auction of multiple items [105]. Bidders submit bids that report their valuations for the items, without knowing the bids of the other bidders. The auction system assigns the items in a socially optimal manner: it charges each individual the harm they cause to other bidders [106]. It gives bidders an incentive to bid their true valuations, by ensuring that the optimal strategy for each bidder is to bid their true valuations of the items. It is a generalization of a Vickrey auction for multiple items. The auction scheme we used is the VCG auction, which requires gathering global information from the system and performing centralized computations. The VCG can achieve a socially optimal allocation [104].

4.1.5 Contributions and Structure of Chapter

A more detailed sharing system has been proposed for the patients to share their edge devices to others nearby. Resolving the competition for finite storage capacity among other patients, we model the patients' actions as a game, which requests the constrained storage capacity from the edge devices. To achieve the optimal results, we formulated the storage request strategies as auction-based strategies and developed the corresponding non-cooperate game mechanism. We assume that patients who request storage are self-interest, and each patient rationally behaves to maximize its own benefit. Patients use weights as a bid for storage capacity and each edge device may assign the extra storage among the patients by itself according to the amount of medical points and channel condition. The patients calculate their bid with proposed game mechanism and send to edge devices. After several round competition, the edge devices choose the winner and cache the medical files by order. In this non-cooperative game, we present Nash Equilibrium as a desirable outcome.

Our analysis concentrates on scenarios with 6, 60, 600, 6000 and 60,000 patients. The results show that the new proposed system can decrease the average transmission delay by 56.87% to 93.69% in the best scenario, which saves more than half time for patients to get medical services. The simulation results also show that the proposed system can decrease the transmission delay by 25.94% to 57.75% even in the worst scenario. In this research:

- A new system for the patients to share their medical records to adjacent edge devices is proposed. Patients can get more efficient health care by caching

EMR on the edge.

- Edge computing is proposed to process and cache EMR. Comparing with traditional cloud computing, edge computing could provide more efficient and safer services. Patient can control their own medical data, made by wearable sensors, on the edge.
- An auction-based and non-cooperative game theory mechanisms is proposed to allocate and cache storage capacity of edge devices. Patients compete the storage resources by considering the four medical factors and communication channel condition.
- We analyzed the performance of proposed system with transmission delay of medical records. The clinical professors can save more than half time to access patients' medical records and provide medical services.

The rest of the chapters are organized as follows. In Section 4.2, we provide the system model with game theory for multiple patients compete edge storage capacity. Section 4.3 focuses on the medical records allocation algorithm and performance analysis. This is followed by the four medical parameters and communication channel situation. In Section 4.4, the system performance evaluation and simulation results have been conducted. Finally, Section 4.5 concludes the chapter.

4.2 System Scenario and Modeling

To explain the system clearly, a case with one "Host Patient" (Patient *A*) and one "Guest Patient" (Patient *1*) has been considered firstly. We consider the patients who share their edge storage as "Host Patient" and the patients who want to compete the storage resources as "Guest Patient". We assume that the registered hospital keeps the whole EMR for each patient. The system is shown in Figure 4.3.

Edge Device *A* and Mobile Device *1* are the edge device of "Host Patient"

(Patient *A*) and "Guest Patient" (Patient *1*) respectively. Two separate edge devices receive real-time medical data from wearable sensors of Patient *A* and Patient *1* respectively, then send it to the nearest hospital after edge processing. Once the real-time data from sensors shows that patient's health condition is not in a normal status. An alert message would be sent by each edge device. The doctors from the nearest hospital would give a timely response and check the patient's health condition. Medical services would be provided if necessary. Edge Device *A* and Mobile Device *1* would cache the medical records that Patient *A* and Patient *1* needed and update the medical data to registered hospital at off-peak time respectively to avoid network congestion, for example at night.

We assume that Patient *A* would like to share his extra capacity storage of edge device with other patients. At the same time, Mobile Device *1* does not have enough storage space for Patient *1* to cache medical records. Mobile Device *1* would send request to Edge Device *A*. Edge Device *A* would response the request and allocate storage capacity to Patient *1*. The Edge Device *A* would update the cached medical data to registered hospital of Patient *1* at off-peak time. In this system, there is no other patients to compete the storage capacity of Edge Caching *A* with Patient *1*. As long as the storage size that Patient *1* requested is smaller than Edge Device *A* provided.

However, there are more patients and edge caches in real life. A system with multiple edge caches and patients has been considered and shown in Figure 4.4. In proposed system, we assume there are there are three "Host Patients" (Patient *A*, Patient *B*, and Patient *C*) would like to share their edge storage and six clusters of "Guest Patients" (Cluster *1*, Cluster *2*, Cluster *3*, Cluster *4*, Cluster *5*, and Cluster *6*) who need extra capacity to cache their medical records. We randomly allocate the position of these six "Guest Patients". The number of each cluster of patients is considered as 6, 60, 600, 6000 and 60,000. However, considering the limited capacity storage of each edge device, competitive relation would be formed by multiple patients in the same game.

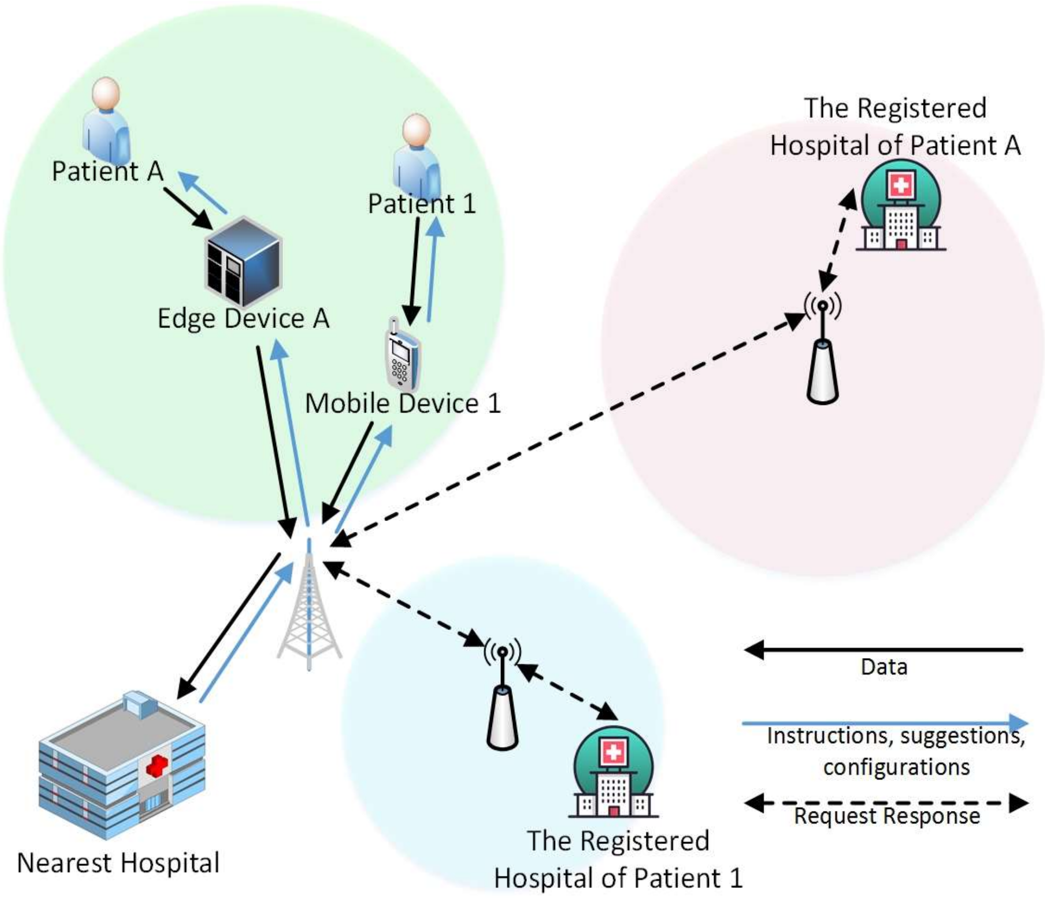

Figure 4.3: The New System with One “Host Patient” (Patient *A*) and One “Guest Patient” (Patient *1*).

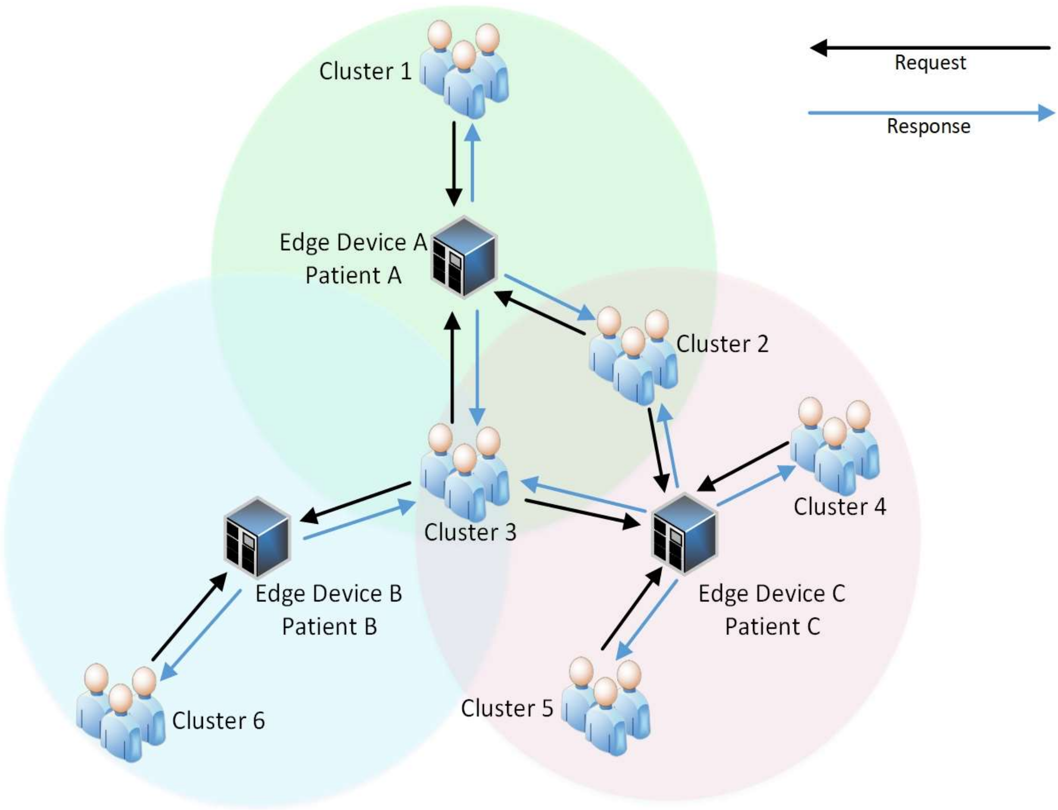

Figure 4.4: Proposed System with Multiple Edge Devices and Patients.

In the proposed system, there are three competitive games in this system. The first auction-based non-cooperative game is to achieve the storage capacity of Edge Caching Device *A*. The candidates of this game are Cluster *1*, Cluster *2*, and Cluster *3*. The second game is to achieve the storage capacity of Edge Caching Device *B*. The candidates of this game are Cluster *3* and Cluster *6*. The third game is to achieve the storage capacity of Edge Caching Device *C*. The candidates of this game are Cluster *2*, Cluster *3*, Cluster *4*, and Cluster *5*. According to the proposed game theory mechanism, patients would know if they are in a high priority position and then select one edge device which provides less waiting time to cache.

The auction-based and non-cooperative game mechanism is explained in next section. The four medical factors and signal channel condition have been considered to compete the winner of each game.

4.3 Mathematical and Analytical Model of Auction-Based Non-Cooperative Game Mechanisms

In this new game mechanism, two parts are considered to compete the edge storage capacity. The first one is the medical parameters of each patient. The second one is channel condition between patients and edge devices. The combination results/weights will be the bid of each patient.

4.3.1 The Medical Factors in Allocation Mechanism

For the medical parameters, four parameters are considered in proposed system.

- The severity of illness (SOI).
- The size of the medical records that patient wants to cache.
- The queueing time that patient has been waited.
- The rewards that patient got by sharing edge caching to other users.

The weights of each medical parameters are showed in following part. The patient who has higher weights has higher priority to cache their medical records in the same edge devices.

a). The severity of illness

The severity of illness (SOI) class [107] is meant to provide a basis for evaluating hospital resource use or to establish patient care guidelines. Patients are assigned their SOI based on their specific diagnoses and procedures performed during their medical encounter. Patients with higher SOI (e.g. major or extreme) are more likely to consume greater healthcare resources than patients with lower SOI in the same diagnosis-related group [108]. The severity of illness has been defined into four levels in the following chart. To compare the priority of patients, we consider the highest level that patient had recently. The patient who had higher level of severity of illness has higher priority to cache their medical records. The weights of SOI (α_n) is presented as in Table 4.3.

Table 4.3: The Weights of Severity of Illness.

LEVEL	Severity	Weights (α_n)
Level 1	Minor	1
Level 2	Moderate	2
Level 3	Major	3
Level 4	Severe	4

b). The size of the medical records that patients want to cache in edge devices

The size of the medical records that patients want to cache is one of the important factors. From the point of sharing storage capacity, the situation that more patients can share the storage capacity on the edge is expected. The patients who want to cache fewer medical records have higher priority. We consider three type of medical records including files/worlds, images, and video. Therefore, there are 7 different combined size need to be considered. The type and weights (λ_n) of the size of medical records are presented in Table 4.4.

c). The queueing time that patients have been waited

The queueing time is considered in the proposed system. After one round of competition, patient could still wait for the same edge device that storage capacity has been allocated to the current winner or change to other edge devices that the patient has higher priority and shorter caching or waiting time. For all the patients in the same game, the one who has waited for a longer time has larger value of weights and higher priority to cache medical files. The weights of the queueing time (β_n) is presented in next table.

Table 4.4: The Type and Size of Medical Records.

The type of medical records	Size of medical records	Weights (λ_n)
Files/words	0G~5G	7
Images	5G~10G	6
Files/words, images	10G~15G	5
Video	15G~20G	4
Files/words, video	20G~25G	3
Images, video	25G~30G	2
Files/words, images, video	30G~35G	1

Table 4.5: The Weights of Queueing Time.

The number of patients	Queueing time	Weights (β_n)
One patient	0~34.13s	1
One patient to two patients	34.13~68.26s	2
One to three patients	68.26~102.39s	3
One to four patients	102.39~136.52s	4
One to five patients	136.52~170.65s	5
One to six patients	170.65~204.78s	6
One to seven patients	204.78~238.91s	7
Other situations	>238.91s	8

d). The rewards that patients got by sharing edge caching to other patients

In the proposed system, patients are encouraged to share their edge storage capacity to other adjacent patients. Edge devices can be shared among other patients to save transmission delay of medical files and improve resources utilization. On the other hand, sharing edge devices means sacrificing devices' battery life and having to deal with a longer processing time. A clear reward algorithm is provided. The patients who share their edge caching to others could get rewards. In this case, the patients who share more to other patients have higher priority to cache in the game. The patient could get one point of reward when sharing their edge to one patient. The weights of the rewards (φ_n) is shown in following table.

Table 4.6: The Rewards System and Weights.

The rewards patient got	Weights (φ_n)
0~20	1
21~40	2
41~60	3
61~80	4
81~100	5
>100	6

The equation for the total weights of four medical factors is shown by (4.1):

$$W_n = \alpha_n + \lambda_n + \beta_n + \varphi_n \tag{4.1}$$

where

W_n: The total weights of four medical factors.

α_n: The weights of severity of illness.

λ_n: The weights of the size of medical records that patients want to cache.

β_n: The weights of queueing time.

φ_n: The weights of the rewards that patients got by sharing edge caching to other patients.

4.3.2 The Channel Capacity between Patients and Edge Devices

For the second part of game theory mechanism, the channel condition between the patients and edge devices has been considered in the system. For the patient who has poor channel condition, it would take too long for transferring their data. Other patients with lower medical weights have to wait for a long time. Therefore, we consider the medical weights selection method with channel condition both. The patients who have higher points in medical weights and better channel condition have higher priority to share the capacity storage of edge devices.

According to the description from the last section, three edge devices have their own communications ranges *r*, individually, which means each edge device is connecting to the patients who are in the corresponding range only as Figure 4.5 shows.

Assuming all patients are clustered into Cluster *1* to *6*, and the number of patients is growing from 6 to 60000, whose positions are initially fixed as an equilateral triangle, therefore, the angle of the equilateral triangle θ = 60°; furthermore, the distance between each two edge devices is *R* and the overlapping distance between the communications ranges of each two edge devices is R_0. Notably, the value of R_0 is shown as below.

$$R_0 = r + r - R = 2 \cdot r - R \tag{4.2}$$

Where in this case, $R_0 > 0$, a parameter η is introduced to describe the ratio between overlapped distance and the communications range, as (4.3) shows,

$$\eta = \frac{R_0}{r} \tag{4.3}$$

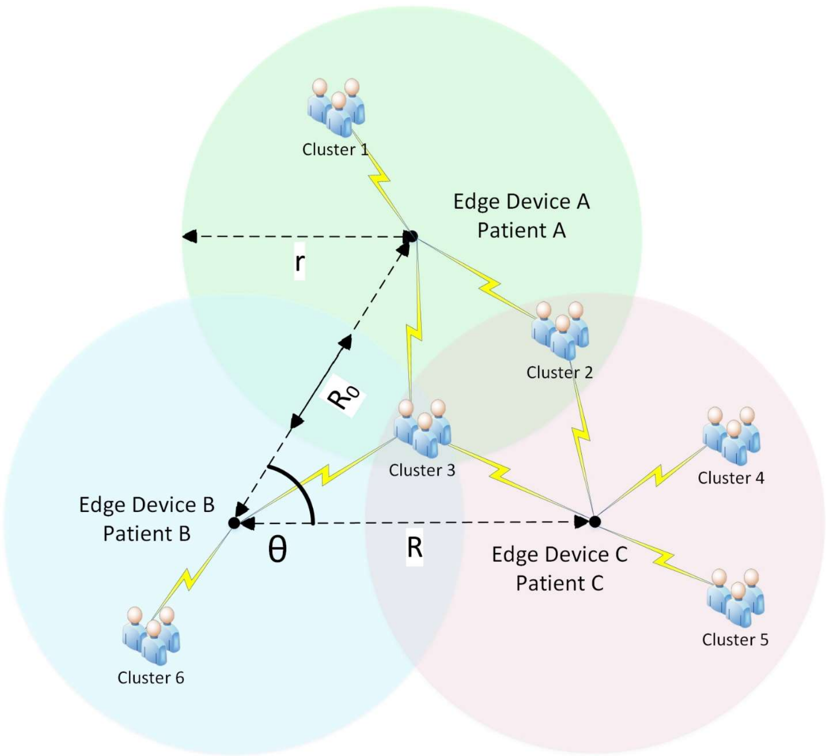

Figure 4.5: Geographical Environment Information for Edge Devices.

Where the value of η is determined by the allocations of mobile patients and the deployment of edge caching devices. Here the value is temporally assumed as 0.75 as Table 4.8 shows, Combing (4.2) and (4.3), the overlapping distance can be yielded as below.

$$R_0 = 2 \cdot r - R = 2 \cdot \frac{R_0}{\eta} - R = R_0 \cdot (\frac{2}{\eta} - 1) \tag{4.4}$$

Where the value of R and R_0 are also determined by the allocations of mobile patients and the deployment of edge caching devices. Here the values are temporally assumed according to Table 4.8.

The simulation of the positions of Patient *1*, Patient *2*, Patient *3*, Patient *4*, Patient *5*, and Patient *6* are following Poisson distribution process in their corresponding ranges.

Therefore, the channel capacity [109] from Patient n to Edge Caching Device q and signal to noise ratio from n to q are R_0 dependent functions. As (4.5) shows,

$$C_{nq}(R_0) = B_{nq} \cdot log_2(1 + \mathrm{SNR}_{nq}(R_0)) \tag{4.5}$$

Here, $C_{nq}(R_0)$ and $SNR_{nq}(R_0)$ are the channel capacity and signal to noise ratio from Patient n to Edge Caching Device q respectively. Accordingly, we have (4.6) as follows,

$$\mathrm{SNR}_{nq}(R_0) = \frac{P_{R_{nq}}}{P_{N_{nq}}} = \frac{P_{T_{nq}} \cdot g_{nq} \cdot d_{nq}(R_0)^{-\varepsilon}}{P_{N_{nq}}} \tag{4.6}$$

Where $P_{T_{nq}}$ is the transmit power; g_{nq} is the transmitter gain; $d_{nq}(R_0)$ is the vector distance; ε is the path loss index; $R_{N_{nq}}$ is the Gaussian white noise [110]. Table 4.8 is the table for simulation parameters. The capacity of communications channel between each patient and its corresponding edge caching device can be simulated.

4.3.3 The Maximum Weights of Patients

The equation by adding the channel condition is updated:

$$\begin{gathered} \mathrm{Max}(R_{nq}) = W_n C_{nq} \\ S_n \leq E_q \end{gathered} \tag{4.7}$$

Where

R_{nq}: The maximum value with medical factors and channel condition.

S_n: The size of medical files for patient n.

E_q: The storage capacity of edge device q.

In (4.6),

$$d_{nq} = \sqrt{\left|x_{T_n} - x_{R_q}\right|^2 + \left|y_{T_n} - y_{R_q}\right|^2} \tag{4.8}$$

Where $x_{T_n}, x_{R_q}, y_{T_n}, y_{R_q}$ is the location of Patient n and Edge Caching Device q.

Therefore, the maximum value can be represented as (4.9):

$$\mathrm{Max}(R_{nq})M = (\alpha_n + \lambda_n + \beta_n + \varphi_n) \cdot B_{nq}\, log_2(1 + \frac{P_{T_{nq}} \cdot g_{nq} \cdot (\sqrt{\left|x_{T_n}(R_0) - x_{R_q}(R_0)\right|^2 + \left|y_{T_n}(R_0) - y_{R_q}(R_0)\right|^2})^{-\varepsilon}}{P_{N_{nq}}})$$

$$S_n \leq E_q \tag{4.9}$$

where

B_{nq}: The bandwidth between Patient *n* and Edge Caching Device q.

g_{nq}: The random variable number accounts for multipath fading follows an exponential distribution with mean $1/\mu$.

$P_{N_{nq}}$: The Gaussian white noise.

The caching time has been considered after all the patient got the value with medical parameters and channel capacity. Especially for the patient who is covered by multiple edge devices. The patient who has the highest combination value of medical parameters and channel capacity in one game doesn't mean the patient has to cache the medical files in corresponding edge devices. Other edge devices could provide better signal connection and faster caching time even though the patient doesn't have the highest value of R_{nq} in this game. Patient has to choose the edge devices with the minimum caching time. Therefore the minimum equation of caching time can be represented as：

$$\mathrm{Min}(T_{nq}) = \frac{M_{nq}}{C_{nq}} + \sum \frac{M_m}{C_{mq}} \tag{4.10}$$

where

T_{nq}: The minimum caching time of Patient *n* in Edge Caching Device q.

M_{nq}: The size of medical files that Patient *n* wants to cache.

C_{nq}: The channel capacity between Patient n and Edge Caching Device q.

M_m: The medical files size of other patients/competitors.

C_{mq}: The channel capacity between other Patients m and Edge Device q.

4.3.4 The Flow Chart of System Model Based on Game Theory

From last three sections, we now understand to add auction-based and non-cooperative game theory to the system model, then comparing and sorting out the priority of diagnostic patients. The algorithm is explained in the Figure 4.6.

Followed by applying (4.9), for the patient who has the highest value and is only covered by one edge device would be cached by corresponding edge device. For the patient who has the highest value and is covered by multiple edge devices, the caching time of the patient in each edge device would be considered. By applying (4.10), the patient would be cached by the edge device which can save the most of time instead of the edge device which has the highest value of R_{nq}. For example, Patient *2* is covered by Edge Device *A* and *C*. In Game *1*, R_{2A} is higher than R_{1A}, Patient *2* has the highest priority to cache. In Game *3*, $R_{4C} > R_{2C} > R_{5C} > R_{3C}$, Patient *2* has the second highest priority to cache. Patient *2* has to compare the caching time by using (4.10), and choose edge devices with the lowest delay to cache. There is a possibility that the caching time of Patient *2* in Game *3* is shorter than Game *1* even though Patient *4* has higher priority than Patient *2*.

The patient who has the highest priority and the shortest caching time would be cached in corresponding edge device. The system would send a request to all the edge devices and ask if more patients need to cache their medical files. If yes, the new game would start again by applying (4.9) and (4.10). If no, the edge device would get the final order for all the covered patients.

According to the flow chart, the medical records transmission delay in the proposed system can be represent in (4.11). In other words, the time that patient can be saved by using proposed system.

$$T(n) = \frac{M_{nq}}{C_{nq}} + \frac{M - M_{nq}}{R} \tag{4.11}$$

where

$T(n)$: The transmission delay of patient n.

M: The maximum size of patient. In proposed System, M is 35G.

R: The transmission rate of macro base station (Bits per seconds).

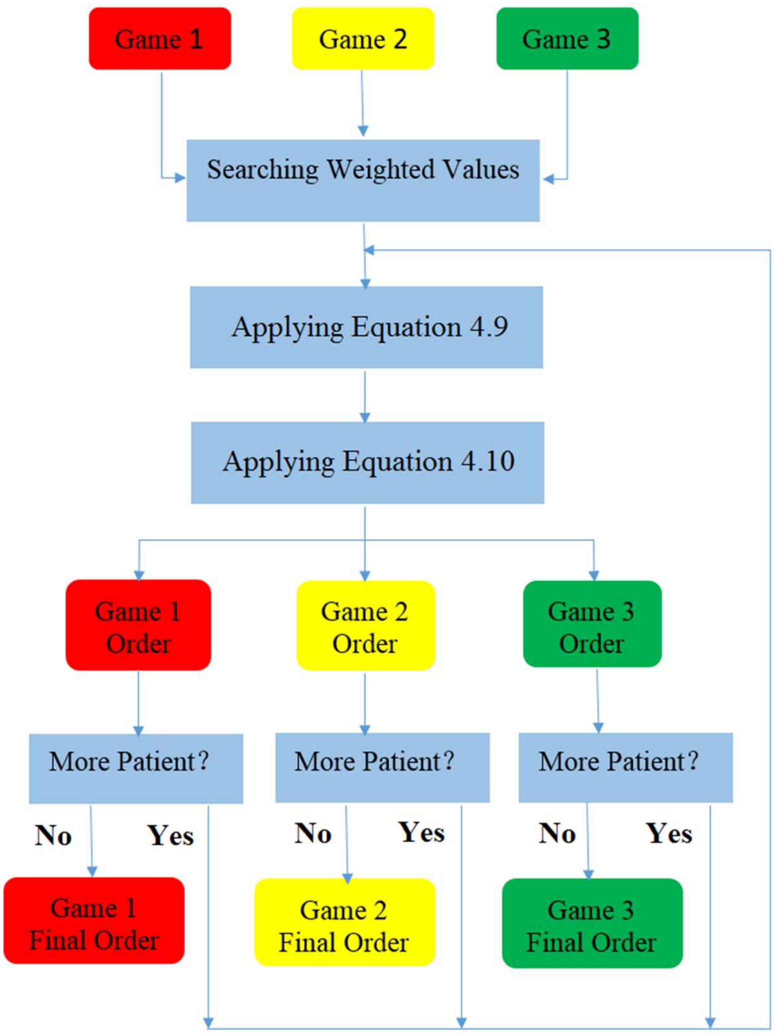

Figure 4.6: Flow Chart of Auction-based Non-Cooperate Game Mechanism.

Table 4.7: The Edge Caching Algorithm.

Algorithm 1: Edge Caching Algorithm

Input:

Patients density in the Study Area λ_N, N=6, 60, 600, 6000, 60000;

Patient q, $q \in N^+$ $(6 \leq N \leq 60000)$;

Edge Caching Device n , $n \in (A, B, C)$;

Overlapping distance, R_0;

Link capacity function, $C_{nq}(R_0)$.

Total weights, $W_n = \alpha_n + \lambda_n + \beta_n + \varphi_n$

Output:

1: Initial simulation of patients' locations,

2: Define patients clusters,

3: **if** *q* drops in Cluster *1* **then**

4: *q* associates to Edge Device *A*, $C_{Aq}(R_0)$,

6: **else**

7: **if** *q* drops in Cluster *4* or *5* **then**

8: *q* associates to Edge Device *C*, $C_{Cq}(R_0)$,

9: **else if**

10: *q* drops in Cluster *6* **then**

11: *q* links to Edge Device *B*, $C_{Bq}(R_0)$,

12: **else**

13: Patient *q* drops in Cluster *2* **then**

14: Patient *q* associates to Edge Device *A and C*, $W_n C_{nq}(R_0)$ where n= *A* and *C*

15: **else**

16: *q* drops in Cluster *3* **then**

17: Patient *q* links to Edge Device *A, B and C*, $W_n C_{nq}(R_0)$ where n= *A, B* and *C*

end

4.3.5 The Other Simulation Parameters of the Studied Scenario

Considering the three scenarios, firstly, the parameter η which describes the ratio between overlapped distance and the communications range is fixed; secondly, the parameter η which describes the ratio between overlapped distance and the communications range is a variable. Thirdly, by applying an auction-based and non-cooperate algorithm, the diagnostic order of all patients will be sorted on bias of weighted value R_{nq}.

Table 4.8: Simulation Parameters for Edge Caching.

Para.	Description	Value
ε	Path loss index	4
Images	5G~10G	6
η	Overlapping ratio R_0/r	0.75
θ	Directional angle between edge devices	60°
$B_{nq}(R_0)$	Effective bandwidth between *n* and *q*	200MHz
$C_{nq}(R_0)$	Link capacity between *n* and *q*	~
g_{nq}	Transmitter gain	2dB
n	Edge caching devices	A,B,C
$R_{N_{nq}}$	The Gaussian white noise	$g_{nq}\sim N(0,N)$
$R_{T_{nq}}$	Transmit power from *n* to *q*	100mW
q	Mobile patients	1,2,3,4,5,6
R	Distance between each two edge devices	300m
r	Communications range	200m
R_0	Overlapping distance	$2 \cdot r - R$
η	Overlapping ratio, R_0/r	0.75

4.4 Performance Evaluation

In this section, firstly, the channel capacity between patients and edge devices has been addressed. Then we combined the medical parameters to calculate the priority of each patient. The simulation results show that the proposed system saves more than half time for clinical professors to provide efficient medical services. The results show that more shared edge devices, higher channel capacity the patients can get, which means the proposed system encourages more patients to share their storage capacity of edge devices to others.

4.4.1 Simulation Results of Channel Capacity

In the proposed system, we considered parameter η is fixed as 0.75. In Game *1*, three clusters of competitors are covered by Edge Device *A*. According to Equation (4.5), the maximum channel capacity between covered clusters and each edge device shows in Figure 4.7, Figure 4.8, and Figure 4.9.

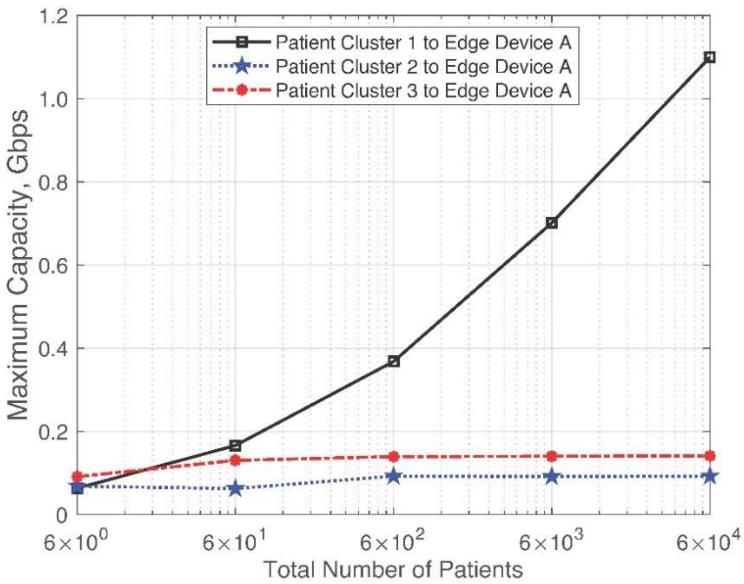

Figure 4.7: Channel Capacity in Game *1* Where All Patients are Covered by Edge Device *A*.

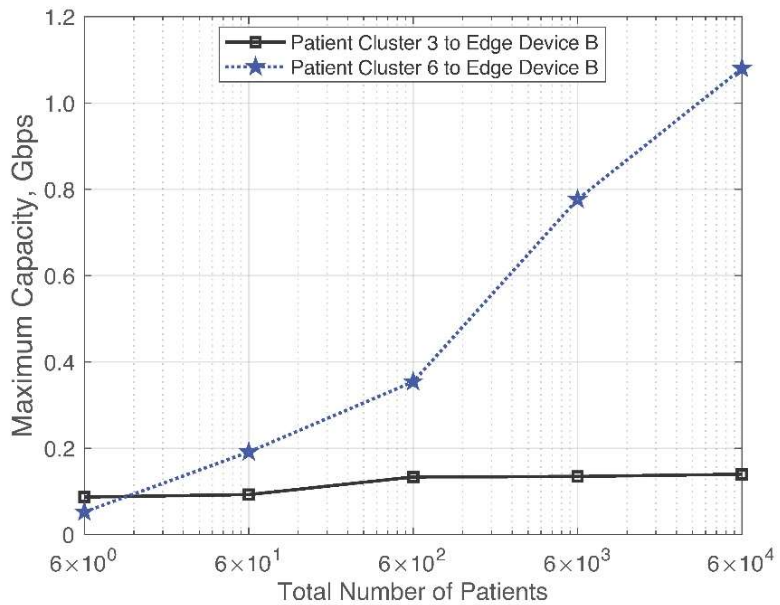

Figure 4.8: Channel Capacity in Game *2* Where All Patients are Connecting to Edge Device *B*.

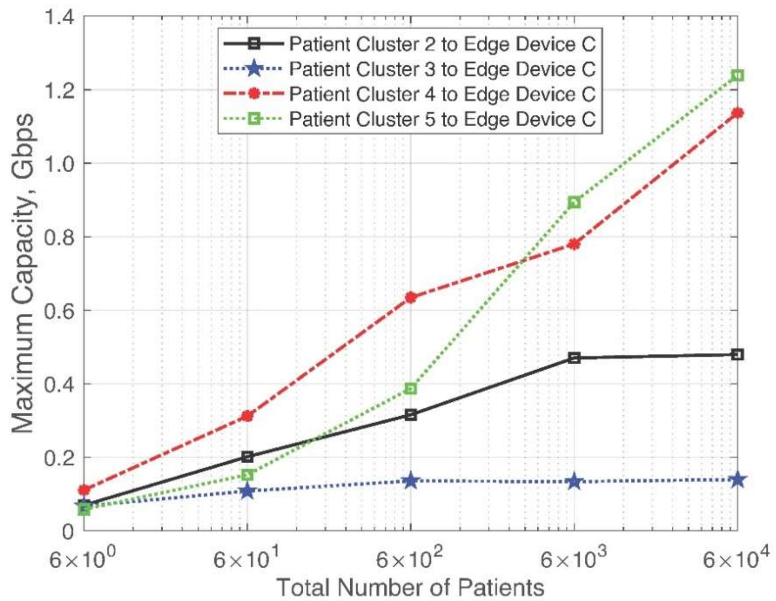

Figure 4.9: Channel Capacity in Game *3* Where All Patients are Connecting to Edge Device *C*.

4.4.2 Maximum value of medical factors and channel condition R_{nq}

In the proposed system, the medical parameters of each cluster of patient are chosen randomly from Table 4.3, 4.4, 4.5, and 4.6. The values of medical weights are showed in Table 4.9.

Followed by using the auction-based and non-cooperative game theory and several round of competitions, the caching order of each game can be simulated and summarized in Table 4.10.

Table 4.9: Medical Weights for All Patients (Unit).

Cluster of Patient	Weighted Values, Max (R_{nq})				
	6	60	600	6000	60,000
1	15	17	14	22	20
2	7	11	13	7	17
3	17	22	10	15	11
4	10	21	18	16	14
5	15	17	14	11	22
6	11	14	18	10	21

Table 4.10: The Allocation of Patients.

Game/Edge	Cluster of Patients			
1/A	Cluster *1*	Cluster *3*	N/A	N/A
2/B	Cluster *6*	N/A	N/A	N/A
3/C	Cluster *2*	Cluster *4*	Cluster *5*	N/A

As we can find from Table 4.10, the medical files of Clusters *1* and *3* are cached by Edge Device *A*. The medical files of Cluster *6* are cached by Edge Device *B*. The medical files of Cluster *4*, Cluster *5*, and Cluster *2* are cached by Edge Device *C* in order.

4.4.3 The Simulation Results of Transmission Delay

The simulation results of transmission delay are compared with the scenario without the proposed system. Two scenarios have been considered: the best situation and the worst situation. The best situation corresponds to the case that the edge devices have enough medical files for the patients. The doctor can provide effective medical services with the cached EMR and does not need to access to whole of EMR which cached in the registered hospital. For the worst situation, it corresponds to the case that the cached EMR in edge devices is not enough for doctor to provide effective and efficient medical care. The doctor need more medical files from registered hospital.

The transmission delay results with the proposed system and without the proposed system in the best situation are shown in Figure 4.10 and Figure 4.11 respectively. As we can find from the figures, the transmission delay has been decreased significantly by applying the proposed system. Especially, for the scenario with 60,000 patients, the transmission delay of Cluster *1*, Cluster *2*, Cluster *3*, Cluster *4*, and Cluster *5* without the proposed system is 512s, 1024s, 512s, 512s, 170.67s, and 170.67s respectively. However, the transmission delay with the proposed system is 13.97s, 60.06s, 108.7s, 13.51s, 4.13s, and 4.74s respectively, which is 97.22% improvement.

For the scenario with 6, 60, 600 and 6000 patients, the system improvement of average transmission delay is 56.87%, 79.85%, 91.19%, and 92.82% respectively. In other words, it can save more than half time to provide medical service for patients.

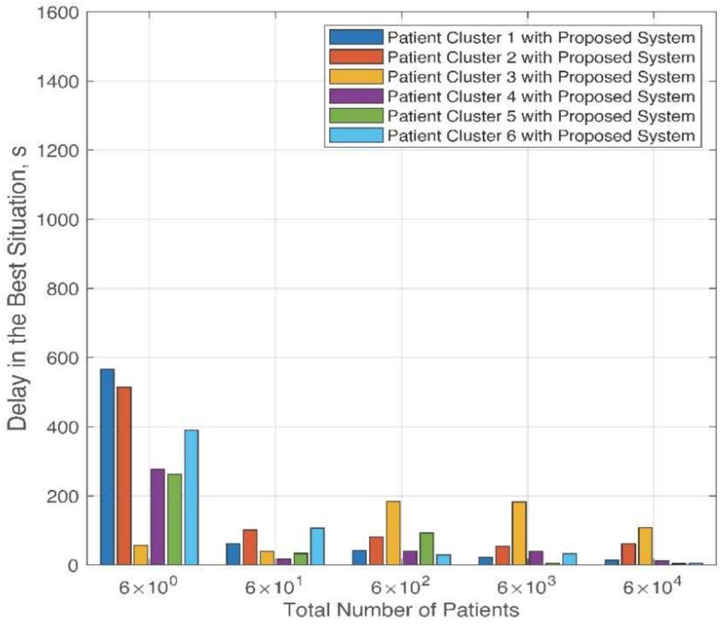

Figure 4.10: The Transmission Delay for All the Patients in the Best Situation.

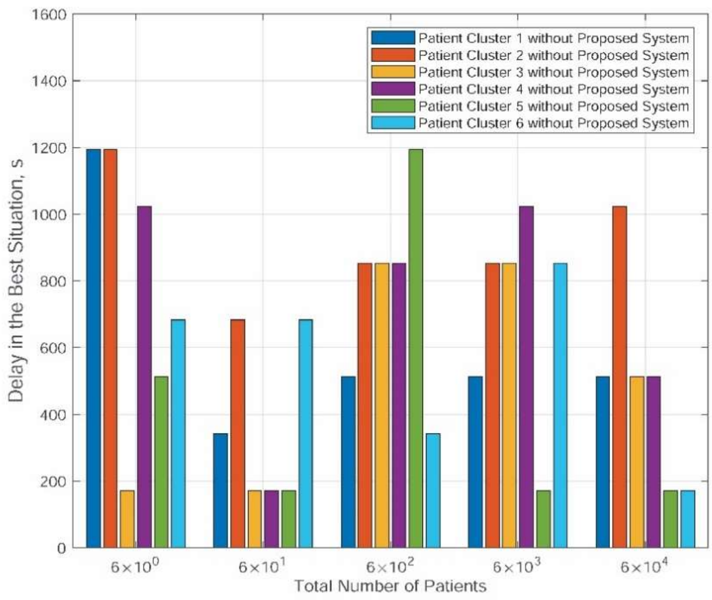

Figure 4.11: The Transmission Delay Without the Proposed System in the Best Situation.

The transmission delay results of the worst situation with the proposed system are shown in Figure 4.12. The transmission delay without the proposed system is 1194.67s for all the patients, which means the whole medical records are required from the registered hospital. The simulation results show that the improvement of transmission delay with proposed system for the scenario with 6, 60, 600, 6000 and 60,000 is 37.85%, 25.94%, 57.75%, 54.82%, 66.71% and 82.9% respectively.

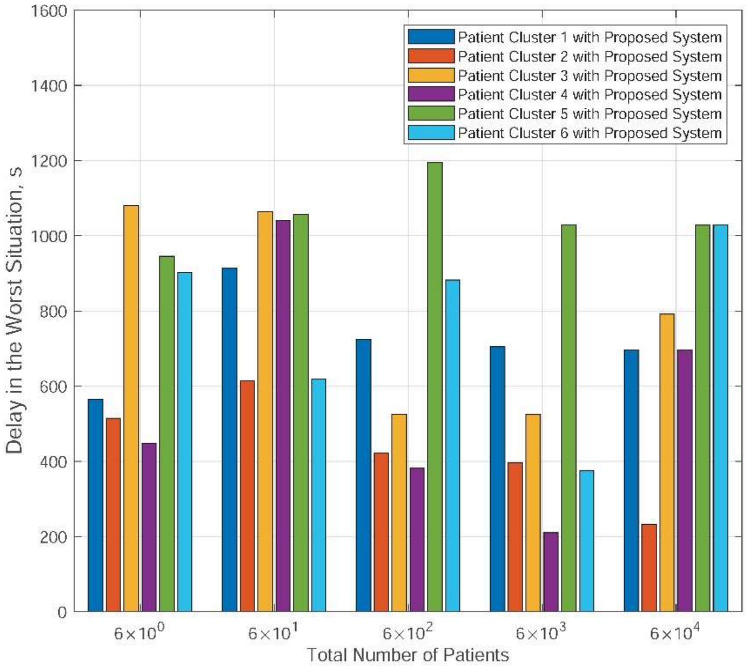

Figure 4.12: The Transmission Delay for All the Patients in the Worst Situation.

4.4.4 The Channel Capacity of Patient When η is changing

In the proposed system with 60000 patients, we assume the overlapping distance is fixed, the data rates each cluster are very close, this is because they are not located in the overlapping areas and singular communications with its corresponding edge caching device is established. However, for Cluster *2* and Cluster *3*, who are seating in the overlapping areas, their connections are usually coordinating with the nearest edge caching device and the data rate of their signal channel are relatively lower. However, when parameter η is not fixed but changing from 0.5 to 1.15, the data rate of each patient is accordantly changed as Figure 4.13 shows.

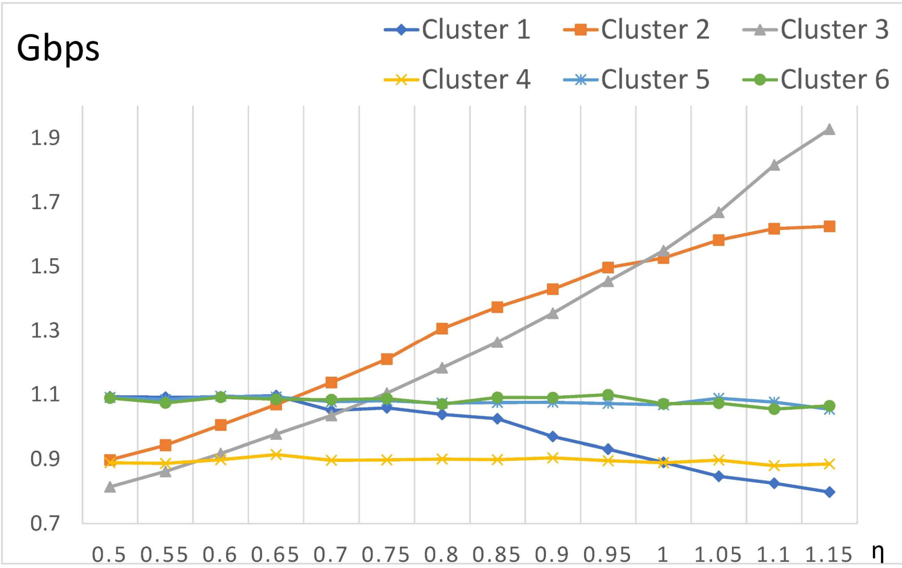

Figure 4.13: Channel Capacity of Each Mobile Patient When η is Changing.

From the Figure 4.13, when the parameter η is increasing, average channel capacity of Cluster *1*, Cluster *4*, Cluster *5*, and Cluster *6* is decreasing; meanwhile, the channel capacity of patients in overlapped are growing significantly, especially for Cluster *3*, which means the more patients share their edge devices, the more communications areas are overlapped, that leads to more choices the patients can have, and the higher channel capacity as a result of utilizing edge caching device.

4.5 Summary and Conclusions

A new system is proposed to share patients' EMR by using edge computing and auction-based non-cooperative game theory mechanisms. We considered the scenario that patients share their storage capacity on the edge device to others and other patients could get effective and efficient medical care by caching their medical files to the adjacent edge devices. In the proposed system, four medical factors including the severity of illness, the size of the medical records that patient wants to cache, the queueing time that patients have been waited, and the rewards that patients got by sharing edge caching with other patients, have been combined with telecommunication channel capacity for patients to compete the allocation and caching priority at the edge. In the proposed non-cooperative game, Nash Equilibrium has been presented as a desirable outcome.

The results show that the proposed system can decrease the transmission delay by 56.87% to 93.69% in the best situation and by 25.94% to 57.75% in the worst situation comparing with the system without sharing edge device. The results show that more shared edge devices, higher channel capacity patients can get, which means the proposed system encourages more patients to share their storage capacity of edge devices to others.

This chapter can give a reference for medical providers to provide more effective and efficient services. They could access patient's medical files in a short time to check patients' current condition with real-time medical data on the edge devices once the emergency happened. They could check the past diagnosis history to give further instructions with the medical files (words, images, and videos) cached at the edge.

Chapter 5

Deep Q-Learning with Preemptive Priority for Sharing EMR

5.1 Introduction

In order to provide efficient and accurate medical care, an important trend in the medical informatics is the use of electronic medical record systems to access to clinical information, provide timely treatment for patients, and prevent the loss or misplacement of medical information [11]. By implementing EMR, multiple healthcare providers can track patient's health condition over a long time [47].

In the last chapter, an auction-based and non-cooperative game theory mechanisms is proposed to allocate and cache storage capacity of edge devices. Patients compete the storage resources by considering the medical parameters and communication channel condition. We analysis the performance of proposed system with transmission delay of medical records. The results show that the proposed system can decrease the transmission delay by 56.87% to 93.69% in the best situation and by 25.94% to 57.75% in the worst situation comparing the system without sharing edge devices. It saves more than half time to access patients' medical files and provide medical services in the best situation. However, in the last system, we considered that all the medical factors including the severity of illness, the size of the medical records that patients want to cache, the queueing time that patients have been waiting, and the rewards that patients got by sharing edge caching with others, have same weights. The priority of allocation is based on a combination value. It is possible that for the patient who has higher illness servility but cannot get timely and effective health care services. On the other hand, in the last system, the medical files of patients can only be cached in a single edge device which could provide the fastest services. However, in our new system, medical records can be cached in multiple edge caching devices dispersedly at same time. The processing time including queueing time and caching time would be decreased significantly. Therefore, we proposed a new sharing system based on deep Q-learning algorithm to allocate and cache the medical records of multiple patients. The learning procedure of the system is to find the optimal Q value for each patient. Q-learning is proposed for patients to find the best routine to cache their medical records among the edge devices. The new system provides more fair and humanized medical care services comparing with the last one. The patients with higher level of illness severity have higher priority and lower waiting time to be served.

5.1.1 Overview of Deep Q-Learning

Deep Q-learning is one of the important algorithms of reinforcement learning. Reinforcement learning is a monitor machine similar to that of the human brain. There is no idea in the head at the beginning. Machine would learn from mistakes through continuous experimentations and find the law and principle and finally achieve the goal [111]. The reinforcement learning agents have achieved some successes in a variety of domains, their applicability has previously been limited to domains in which useful features can be handcrafted, or to domains with fully observed, low-dimensional state spaces. A novel artificial agent, termed a deep Q-learning [29] is developed by using recent advances in training deep neural networks, which can learn successful policies directly from high-dimensional sensory inputs using end-to-end reinforcement learning. Deep Q-learning was proposed in [112] primitively. Deep Q network is a method that combines Q-Learning and neural network. Before we study deep Q network, the introduction of Q-learning and neural network are studied in the following part.

a) What is Q-Learning?

We all have our own code of conduct when we start doing something. For example, when we were young, parents said to us that you cannot watch TV without finishing homework. Hence in the state of writing homework, the good behavior is to continue writing homework until we finished. After finishing it, we can get rewards. The bad behavior is to watch TV without writing homework. As a result, patients would give seriously criticism if they found out. It became our indelible memory because it happens too many times. This is also the code of conduct of Q-learning that we want. Q-learning is a decision-making process.

For example, we are in the state of writing homework and we have never tried to watch TV while writing homework. Therefore, we have two choices under this state. The first one is to continue writing homework. The second one is to watch TV. Because I have not been punished before, I choose to watch TV, and now the state has been changed to watch TV. Then I choose to continue watching TV. Finally, parents went home and found out that I went to watch TV without finish homework. They punished me, and I also deeply recorded this experience, which changed the behavior of "watching TV without finishing homework" into negative or bad behavior in my mind.

Assume that we have learned code of conduct. Now we are in state of writing homework (*s1*). I have two behaviors or actions including watching TV (*a1*) and writing homework (*a2*). According to the experience, under the state of *s1*, the potential rewards of *a2* are higher than *a1*. Then the potential rewards can be replaced by a Q table with states and actions. In the Q table, the reward of *Q (s1, a1)* is less than *Q (s1, a2)*, therefore *a2* is chosen as the next behavior. Now the state is updated to *s2*. Under the state of *s2*, we still have two identical choices and repeat the above process. According to the code of conduct in Q table, we compare the rewards of *Q (s2, a1)* and *Q (s2, a2)* and choose the larger one. Then we will reach to state *s3* according to *a2* and repeat the above decision process. This is the algorithm that how Q-learning makes decisions.

Table 5.1. Q Table with States and Actions.

Q Table	*a1*	*a2*
s1	-2	1
s2	-4	2

We go back to the previous process, according to the estimation of the Q table, the value of *a2* is larger in *a1*. By applying the previous decision-making method, we took *a2* in *s1* and reached *s2*, then we start to update Q table for decision-making. For *s2*, we did not take any action in practice, but we imagine that we take an action from *a1* and *a2*, and find out which action has larger Q value. If the value of *Q (s2, a2)* is larger than *Q (s2, a1)*, then we multiply *Q (s2, a2)* by an attenuation value gamma (for example, 0.9) and plus the reward *R* that we obtained when it reaches *s2*, which is considered as the value of *Q (s1, a2)* in practice. However, we used estimated value of *Q (s1, a2)* according to the Q table. We can now update *Q (s1, a2)* based on the difference between the value of reality and estimation. We multiply the gap by a learning efficiency parameter alpha and plus the previous value of *Q (s1, a2)*, then becomes a new and updated value. This is the optimization process that how Q-learning algorithm updates Q table and makes decisions.

b) What is Neural Network?

Neural network is a computational model that is inspired by the way that biological neural networks in the human brain process information. Neural networks are a set of algorithms that are designed to recognize patterns. They interpret sensory data through a kind of machine perception, labelling or clustering raw input. The patterns they recognize are numerical, contained in vectors, into which all real-world data, be it images, sound, text or time series, must be translated. For example, in image recognition, they might learn how to identify images that contain cats by analysing example images that have been manually labeled as "cat" or "no cat" and using the results to identify cats in other images. They do this without any prior knowledge about cats, for example, that they have fur, tails, whiskers and cat-like faces. Instead, they automatically generate identifying characteristics from the learning material that they process. Artificial neural networks have generated a lot of excitement in machine learning research and industry, thanks to many breakthrough results in speech recognition, computer vision and text processing. The neural network can be systematically divided into CNN (Convolutional Neural Network) with spatially distributed data and RNN (Recurrent Neural Network) with time-distributed data.

CNN is a class of deep neural networks, most commonly applied to analysing visual imagery. A CNN consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of a series of convolutional layers that convolve with a multiplication or other Iot products. For example, the process of face recognition by applying CNN. Based on the CNN training of the first layer, partial and independent modules of image can be recognized, such as a square, a circle, or an eye. A large amount of image data is inputted at this layer, the neurons of this layer have the capability to distinguish the boundaries of graphics. The following layers would recognize a higher level of image mode (for example, the combination mode with a few small square and circle) based on the information obtained in the previous layer. Then a face can be recognized by a quick manner.

RNN is a class of artificial neural networks where connections between nodes form a directed graph along a temporal sequence. This allows it to exhibit temporal dynamic behaviour. Unlike feedforward neural networks, RNNs can use their internal state (memory) to process sequences of inputs. This makes them applicable to the tasks such as unsegmented, connected handwriting recognition, natural language processing, or speech recognition. Language is time-distributed data, which means that the meaning of the next sentence is related to the previous sentence. The important characteristic of the RNN is to remember previous information. For example, the following conversation:

"How much data does the 10 pounds package include?"

"2GB internet data is included."

"What about 20 pounds?"

This conversation is easy to understand for us. The last question is obviously asking "How much data does the 20 pounds package include?". However, if we want machine to understand this conversation, the previous package information is needed, which requires a loop design that allows the neural network to have a depth in time.

c) Deep Q-Learning

As we mentioned, deep Q-learning is a combination algorithm of Q-learning and neural network. In Q-learning we use tables to store each state, however there are millions of states in practice. If we use tables to store, the computer or machine has not enough storage capacity. It costs huge efforts and time to search for the corresponding states every time in such a large table. However, in machine learning, neural network is good at dealing with complex states. States and actions are used as input to the neural network, and then we can get Q value of the action through analysis of neural network. By applying neural network, we do not need to record the Q value in the table, instead generating the Q value directly. We can imagine that the neural network accepts the external information, which is equivalent to collect information from our eyes and ears, and then process in brain and output the value of each action, and finally select one of the actions with better Q value by reinforcement learning.

Deep Q-learning is considered tasks in which the agent interacts with an environment through a sequence of states, actions and rewards [30]. The goal of the agent is to select actions in a fashion that maximizes the cumulative future rewards. More formally, we use a deep Q-learning to approximate the optimal action-value function (Q values).

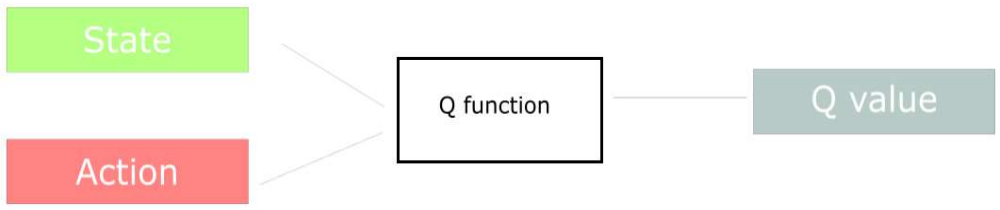

Figure 5.1: Deep Learning with Q Function.

The action value function (or "Q-function") takes two inputs: "state" and "action". It returns the expected future reward of that action at that state. The equation for Q value is shown as following [113].

$$Q^{\pi}(s_t, a_t) = \text{Max}\, E\, [R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \ldots | s_t, a_t] \qquad (5.1)$$

Where

Q: The maximum sum of rewards with given state and action.

π: The behavior policy.

s_t: The given state or current state.

a_t: The given state or current action.

E: The estimated discounted cumulative rewards with future rewards.

R_{t+1}, R_{t+2}, R_{t+3}: The future rewards.

γ: The attenuation parameters of future rewards.

From the equation, the Q value is the sum of all future rewards, but these rewards are attenuating. When γ is 1, the Q value is a reward without attenuation in the future, which means the full value of all the steps. When γ is 0, the Q value is a reward which is concerned with current reward. If the value of γ changes from 0 to 1, which means the rewards are not only related to current interest, but also the interests for the future.

The Q-learning algorithm process is shown in the Figure 5.2. The initial Q table chooses and executes an action. The Q function would measure the Q values based on the action. The learning system would stop updating the Q value if the optimised Q table is found.

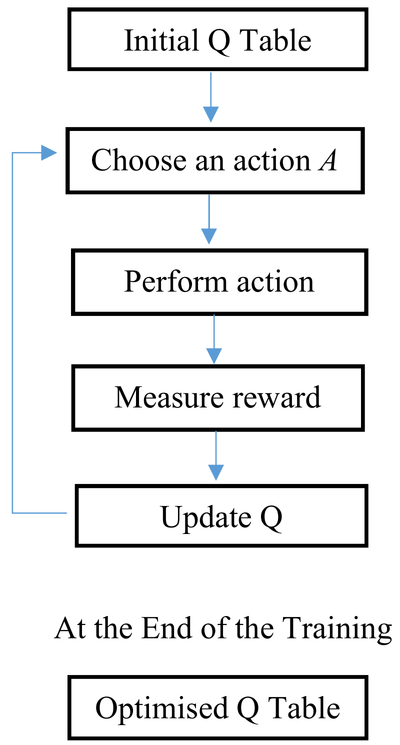

Figure 5.2: The Q-Learning Algorithm Process.

5.1.2 Queueing Model with Preemptive Priority

Much of queueing theory is devoted to analysis priority queues, where customers are labelled and served in accordance with a priority scheme, high-priority customers are served with less waiting time and low-priority customers in the queue. Compared with the general queuing model, many practical queuing systems are more adaptive with the queuing theory model with priority, such as the healthcare area. The patients with higher priority should be served firstly comparing with other patients. Therefore, the queuing theory model with priority has more practical significance. Queueing model can be divided into non-preemptive priority and preemptive priority [114].

For the non-preemptive priority, which indicates that even if a high-priority user arrives, the system cannot stop a low-priority user who is receiving service to return to the queue. In other words, once the user is served at moment, the service cannot be interrupted until this user finished.

For the preemptive priority, which means that once a high-priority user arrives, the system immediately interrupts the service to the low-priority customer (return to the queue) and immediately starts to serve high-priority users.

5.1.3 Contributions and Structure of Chapter

More formally, we use a deep Q network to approximate the optimal action-value function (Q values). In this chapter, we consider four states in the allocation algorithm including the illness servility of patients, the storage capacity of edge devices, the processing time of medical files and the channel condition between the patients and edge devices. The action of this learning algorithm is the caching decision (1/0 or yes/no) of edge devices for each patient. The learning procedure of the system is to find the optimal Q value for each patient. Q-learning is proposed for patients to find the best routine to cache their medical records among the edge devices. The results show the transmission delay could at least decrease the average transmission delay by 16.78% and 71.14% comparing the scenario with the previous system and without applying the proposed system. The total transmission time is analysed as well, the results show that new allocation algorithm could save at least 20.57% and 60.08% time for the whole of system comparing with the previous system and the scenario without proposed algorithm. Therefore, in this chapter:

- The illness servility of patient has been considered at the first place, the patient who has higher illness servility has higher priority to cache.
- The medical files of patients can be cached in multiple edge caching devices dispersedly instead of one edge caching. The processing time including queueing time and caching time would be decreased significantly.
- A new reward algorithm has been proposed to decide priority of patients.
- The queueing time model based on preemptive priorities (M/M/1) has been proposed to calculate the processing time.
- Deep Q-learning has been proposed in edge devices to find the best caching routine for patients.

The rest of the chapter is organized as follows. In Section 5.2, we provide the system model with deep Q-learning based on four states. Section 5.3 focuses on the medical records' allocation method and performance analysis. The simulation results have been conducted. Finally, Section 5.4 concludes the chapter.

5.2 System Scenario and Modeling

Three parts are considered in this section. The first one shows the four states of learning system. The second part is proposed to find the best Q value and allocate medical records to the edge devices. The model of transmission delay is proposed in the third part.

For the deep Q-learning algorithm, four parameters are considered in proposed system, which include the severity of illness (SOI), the storage capacity of edge devices, the processing time of medical files and the channel condition between patients and edge devices. The system model of each state is showed in following part. The patient who has higher Q value has higher priority to cache their medical records.

5.2.1 The Four states in Deep Q-Learning

1) The illness servility

The severity of illness (SOI) class is meant to provide a basis for evaluating hospital resource use or to establish patient care guidelines. Patients are assigned their SOI based on their specific diagnoses and procedures performed during their medical encounter. Patients with higher SOI (e.g. major or severe) are more likely to consume greater healthcare resources than patients with lower SOI in the same diagnosis-related group.

Table 5.2: The Weights of Severity of Illness.

LEVEL	Severity	Rewards $R_I(n)$
Level 1	Minor	e^1
Level 2	Moderate	e^2
Level 3	Major	e^3
Level 4	Severe	e^4

The severity of illness has been defined into four levels. For the level 1, it means that the patients need non-invasive diagnostic or minor therapeutic. For the moderate and major level, which mean that patients need therapeutic and invasive diagnostic, and non-emergency life support respectively. The highest level, severe, which means the patients need emergency life support immediately. To compare the caching priority of patients, we consider the highest level that patient had recently. The patient who had high level of severity of illness, for example, major or severe, are more likely to have the same or higher level again. Therefore, they have higher priority to cache their medical records firstly. The rewards of illness servility ($R_I(n)$) is proposed to apply Exponential Function. The weights of SOI (α_n) is presented as in Table 5.2.

2) The storage capacity of edge devices

Different with the last research, in this proposed system, the patients can cache their medical files to different edge devices instead of caching all in one edge device. The caching algorithm is based on queueing time and channel condition of each edge device. An example with two patients and two edge devices is explained, which is shown in Figure 5.3.

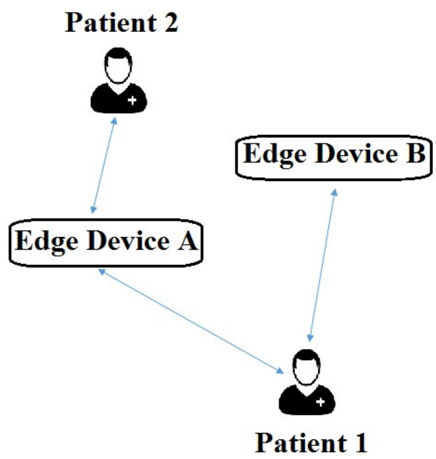

Figure 5.3: A Scenario for Two Patients and Two Edge Devices.

Patient *A* could request storage capacity from Edge Devices *A* and *B*. Patient *B* requests storage capacity from Edge Device *A*. We assume that Patient *2* has higher illness severity than Patient *1*. The channel capacity between Patient *1* to Edge Device *A* is better than Edge Device *B*. In this system, Patient *2* has the highest priority in Edge Device *A* and starts to cache. At the same time, Patient *A* starts to cache medical files in Edge Device *B*. Once Patient *2* cached medical files in Edge Device *A*, three questions would be asked:

"Dose Patient *2*/Higher illness servility patients finish their caching?"

"Do I (Edge Device *A*) still have enough space to share?"

"Is Patient *1* still caching or processing his/her medical files?"

If all the answer is yes for all three questions. The Patients *1* would change to Edge Devices A to cache the rest of medical files, which not only decreasing the processing delay, but also saving more time for other patients who want to cache.

3) The processing time of medical files

In proposed system, the processing time includes the queueing time and caching time. For the queuing model, we proposed the model with priority, the order in which members of the team are served is based on the priority they are given. We applied M/M/1 queueing model with preemptive priority for processing time. In the system, preemptive priority follows two rules: The first one is that once a high priority patient arrives, the edge device immediately interrupts the services to the low priority patient (This patient would return to the queue or choose other edge devices) and starts to serve the high priority patient immediately. After the caching of medical files is finished, the edge device would follow the caching algorithm and choose the next patient to be served. The second is that for the patients with same priority, the system follows the algorithm of "First come, first serve". Patients would not be served until the services of other patients with same priority are done. Therefore, the equation of the processing time is shown by (5.2):

$$T_{nq}(k) = \frac{1/\mu_q}{\left(1 - \frac{\sum_{i=1}^{k} \lambda_{i-1}}{G_n \mu_q}\right)\left(1 - \frac{\sum_{i=1}^{k} \lambda_i}{G_n \mu_q}\right)}$$

$$\text{For } k = 1, 2, 3, \dots, K \tag{5.2}$$

According to the little equation, the equation of queueing time can be presented by (5.3):

$$W_{nq}(k) = T_{nq}(k) - 1/\mu_q \tag{5.3}$$

where

$T_n(k)$: The processing time for patient n with priority k.

μ_m: The mean service rate per edge device.

λ_i: The mean number of patients who want to cache for priority i.

G_n: The number of edge devices that can serve Patent n.

$W_n(k)$: The queueing time for Patient n with priority k.

4) The channel capacity

The channel capacity is considered as one of the states. For the edge devices that patient has same queueing time, patients would choose the edge devices with higher channel capacity to save more caching delay. Thus, the channel with higher capacity should have higher Q values. The channel capacity from Patient n to Edge Caching Device q and signal to noise ratio from n to q as (5.4) shows,

$$C_{nq}(d) = B_{nq} \cdot log_2(1 + \mathrm{SNR}_{nq}(d)) \tag{5.4}$$

Here, $C_{nq}(d)$ and $SNR_{nq}(d)$ are the channel capacity and signal to noise ratio from Patient n to Edge Caching Device q respectively. Accordingly, we have (5.5) as follows,

$$\mathrm{SNR}_{nq}(d) = \frac{P_{R_{nq}}}{P_{N_{nq}}} = \frac{P_{T_{nq}} \cdot g_{nq} \cdot D_{nq}(\mathrm{d})^{-\varepsilon}}{P_{N_{nq}}} \tag{5.5}$$

Where $P_{T_{nq}}$ is the transmit power, g_{nq} is the transmitter gain, $D_{nq}(\mathrm{d})$ is the vector distance, d is the distance between the patient and edge devices, ε is the path loss index, $R_{N_{nq}}$ is the Gaussian white noise. The capacity of communications channel between each patient and its corresponding edge caching device can be simulated.

5.2.2 Allocate and Cache Medical Records with the Best Q values

From the last three sections, we now understand to add deep Q-learning to the system model. The system would compare and sort out the priority of diagnostic patients according to the Q values. The allocation algorithm is explained in a flow chart, which is shown in Figure 5.4.

We assume Patient *1* wants to cache medical files in Edge Device *A*. Before caching medical files, Patient *1* would send a request to Edge Device *A* and check if any other patients are caching now. If yes, then Patient *A* has to compare the priority with other patients. The edge device would stop processing current patient and serve Patient *1* only when Patient *1* has higher priority than others. Then Patient *1* would ask if there is enough storage capacity in Edge Device *A*. If no, Patient *1* changes to other edge devices. If yes, Patient *1* has to compare the illness severity level with other patients who want to cache their medical files in Edge Device *A* as well. If Patient *1* has the highest illness severity and the shortest processing time, then the medical files would be cached in Edge Devices *A*. During the caching, if other edge caches could offer a less processing time, then Patient *1* would change to other caches until all the medical files have been allocated. So according to the flow chart, the maximum rewards (Q values) can be presented by (5.6).

$$\mathrm{Max}\left(\sum_{q} R_{nq}\right) = e_n^{\alpha} \cdot \frac{1}{\left(\sum_{q=1}^{m} \frac{S_{nq}}{\sum_{q=1}^{m} C_{nq}} + W_{nq}(k)\right)} \cdot C_{nq}(d) - \left(\frac{\sum_{q=1}^{m} T_{nq}(k)}{m} - T_{nQ}(k)\right)$$

$$SE_q > 0 \tag{5.6}$$

where

$T_{nQ}(k)$: The minimum processing time of Patient n.

The ideal situation is that the practical processing time equals to the minimum processing time.

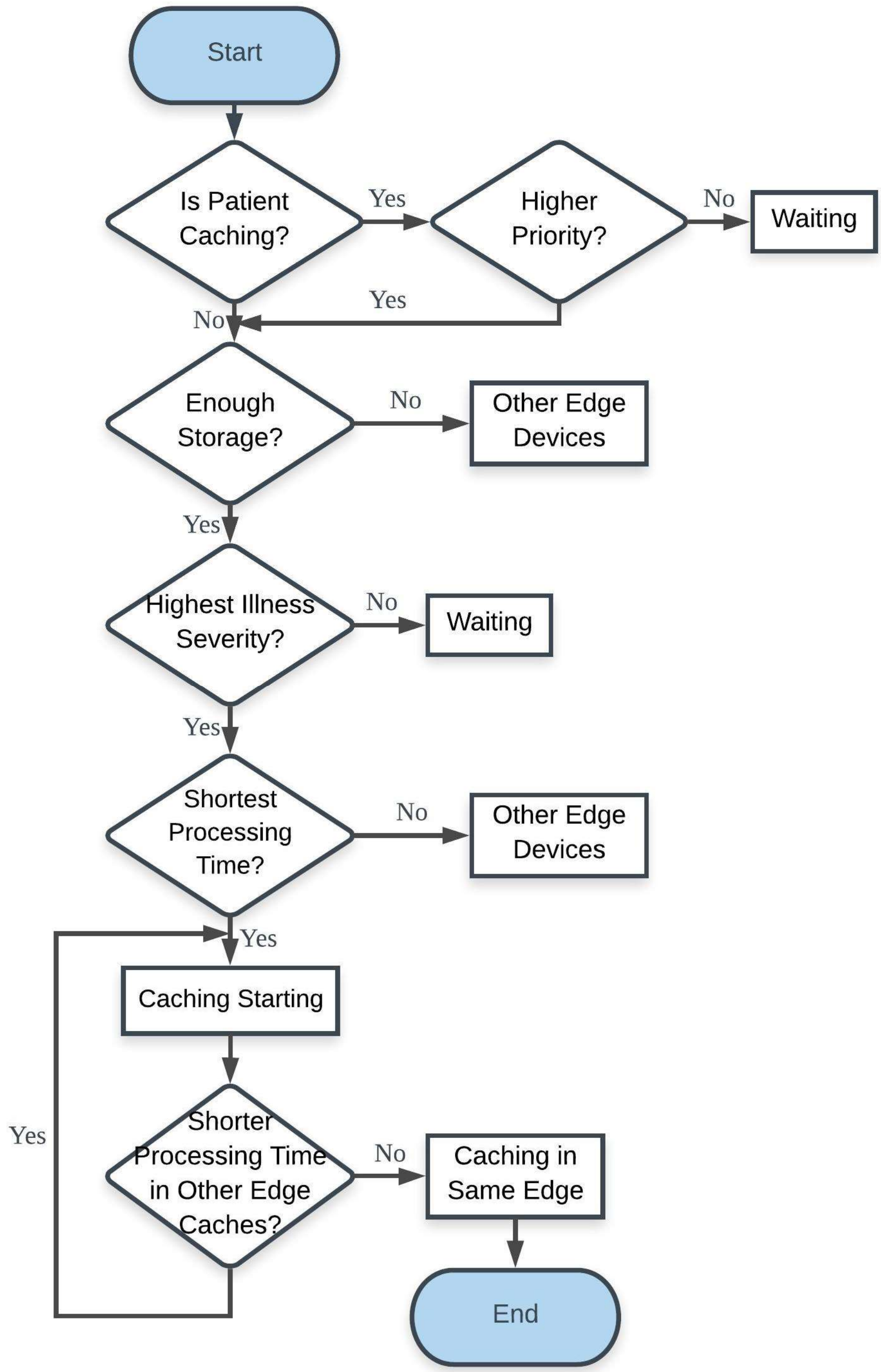

Figure 5.4: The Flow Chart of Allocation Algorithm Based on Deep Q-Learning.

5.2.3 The Transmission Delay with Proposed System

The caching time has been considered after all the patient got the value with medical parameters and channel capacity. Especially for the patient who is covered by multiple edge devices. Patients choose the edge devices with the minimum queueing time. Therefore, the equation of caching time can be represented as：

$$D_n = \sum_{q=1}^{m} \frac{S_{nq}}{\sum_{q=1}^{m} C_{nq}} + \frac{S_{j1}}{C_{j1}} \tag{5.7}$$

Where S_{j1}/C_{j1} is the queueing time for the first edge caching device.

5.3 Performance Evaluation

We applied new allocation system based on deep Q-learning in our last scenario. In the last system, we assume that there are three "Host Patients" (Patient *A*, Patient *B*, and Patient *C*) would like to share their edge storage to six clusters of "Guest Patients" (Cluster *1*, Cluster *2*, Cluster *3*, Cluster *4*, Cluster *5*, and Cluster *6*) who need extra capacity to cache their medical records, which shows in Figure 5.5 We randomly allocate the position of these six clusters of "Guest Patients". However, considering the limited capacity storage of each edge device, competitive relation would be formed by multiple patients in the same game.

In proposed system, there are three competitive games in this system. The first auction-based non-cooperative game is to achieve the storage capacity of Edge Caching Device *A*. The candidates of this game are Cluster *1*, Cluster *2* and Cluster *3*. The second game is to achieve the storage capacity of Edge Caching Device *B*. The candidates of this game are Cluster *3* and Cluster *6*. The third game is to achieve the storage capacity of Edge Caching Device *C*. The candidates of this game are Cluster *2*, Cluster *3*, Cluster *4* and Cluster *5*. According to the proposed game theory mechanisms, patients would know if they are in a high priority position and then select one edge device which provide less waiting time to cache.

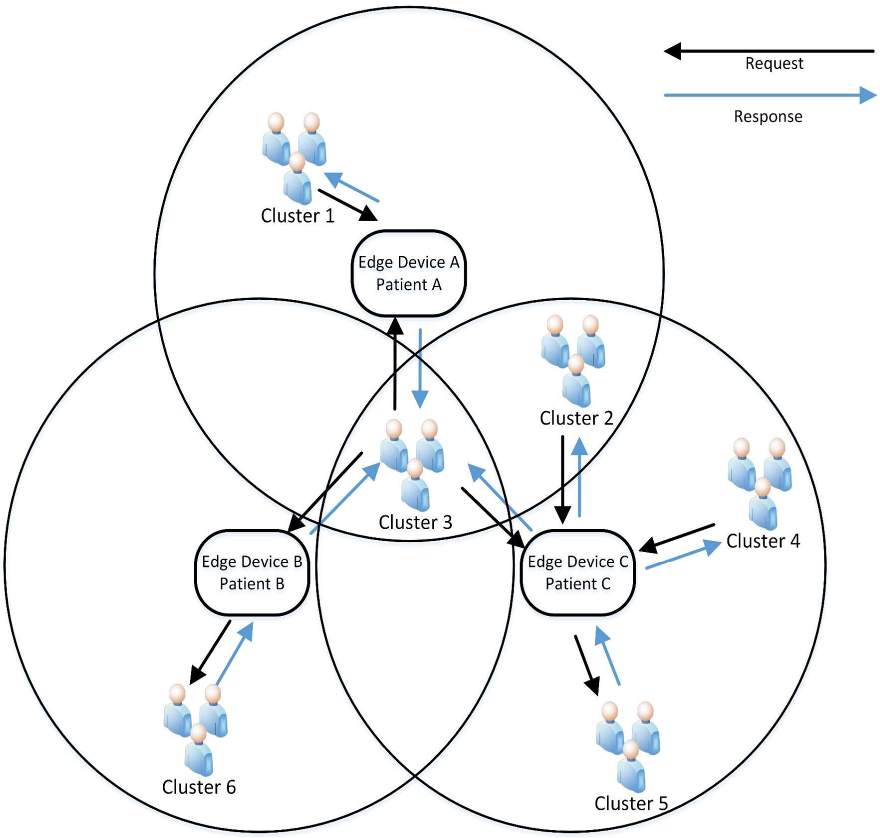

Figure 5.5: A Scenario with "Host Patients" and "Guest Patients".

Table 5.3: Patients Allocation and Caching Order.

Game	Caching order			
1	Cluster *1* (100%)	N/A	N/A	N/A
2	Cluster *6* (100%)	Cluster *3* (100%)	N/A	N/A
3	Cluster *4* (100%)	Cluster *5* (100%)	Cluster *2* (100%)	N/A

As we can find from Table 5.3, the medical files of Cluster *1* are cached by Edge Device *A*. The medical files of Cluster *3* and Cluster *6* are cached by Edge Device *B*. Cluster *3* has the highest channel capacity with Edge Device *A* comparing with Edge Device *B* and *C*. However, in Edge Device *B*, Cluster *3* costs less caching time due to Cluster *6* has a high channel capacity. The medical files of Cluster *4*, Cluster *5*, and Cluster *2* are cached by Edge Device *C* in order. In proposed deep Q-learning system, each patient has to find the edge device with the best Q value and other edge devices that are available to cache. According to the flow chart and Equation (5.6), the new allocation results are shown in Table 5.4.

Table 5.4: Patients Allocation and Caching Order in Proposed System.

Edge	Caching order			
A	Cluster *1* (100%)	Cluster *2* (23.1%)	N/A	N/A
B	Cluster *3* (50.02%)	Cluster *6* (100%)	N/A	N/A
C	Cluster *3* (49.98%)	Cluster *4* (100%)	Cluster *2* (76.9%)	Cluster *5* (100%)

Different with last research, in proposed new system, the medical files of Cluster *1* and 23.1% medical files of Cluster *2* are cached in Edge Device *A*. The left medical files of Cluster *2* are cached in Edge Device *C*. For the Cluster *3*, 50.02% and 49.98% medical files are cached in Edge Device *B* and *C* respectively. The medical files of Cluster *4* and *5* are cached in Edge Device *C*. For the Cluster *6*, the medical files are cached in Edge Device *B*. The caching order of scenario shows in Figure 5.6.

According to the Equation (5.6), the transmission delay results of proposed system based on deep Q-learning are shown in Figures 5.7, 5.8, and 5.9.

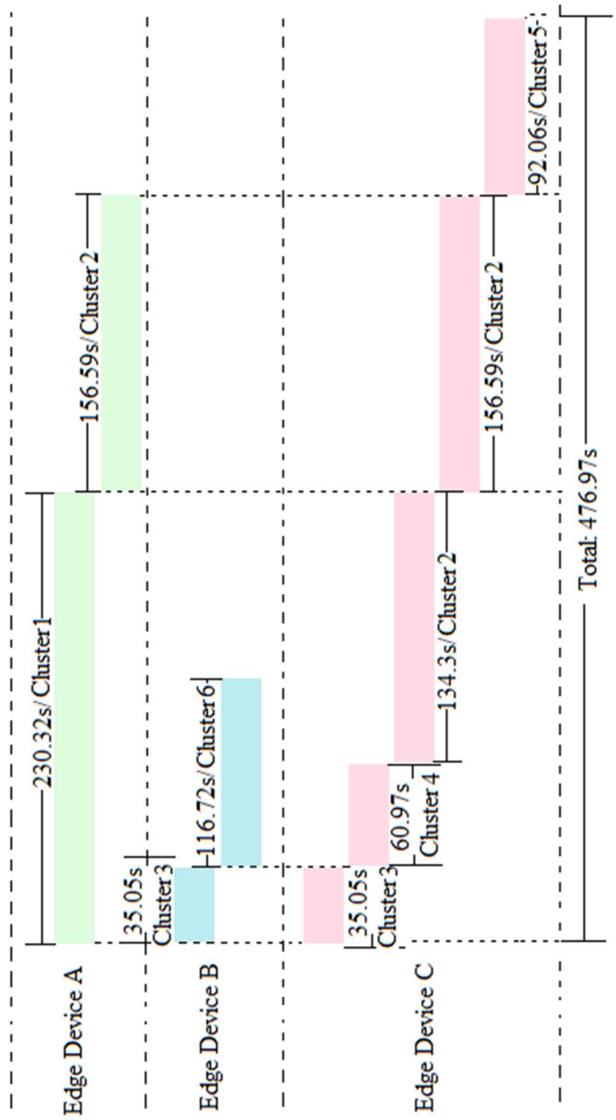

Figure 5.6: The Allocation and Caching Results of Patients in Each Edge Device.

For Cluster *1*, Cluster *2*, Cluster *3*, Cluster *4*, Cluster *5*, and Cluster *6*, the transmission delay with proposed system is 230.32s, 386.91s, 35.05s, 96.02s, 478.97s and 151.77s respectively. However, by applying previous system, the transmission delay is 230.32s, 508.45s, 70.1s, 60.97s, 600.51s, and 186.82s respectively. The transmission delay without proposed system is 1194.67s, 1194.67s, 170.61s, 1024s, 512s, and 682s respectively, which is an 80.72%, 75.65%%, 82.39%, 90.62%, 25.19%, and 77.75% improvement respectively.

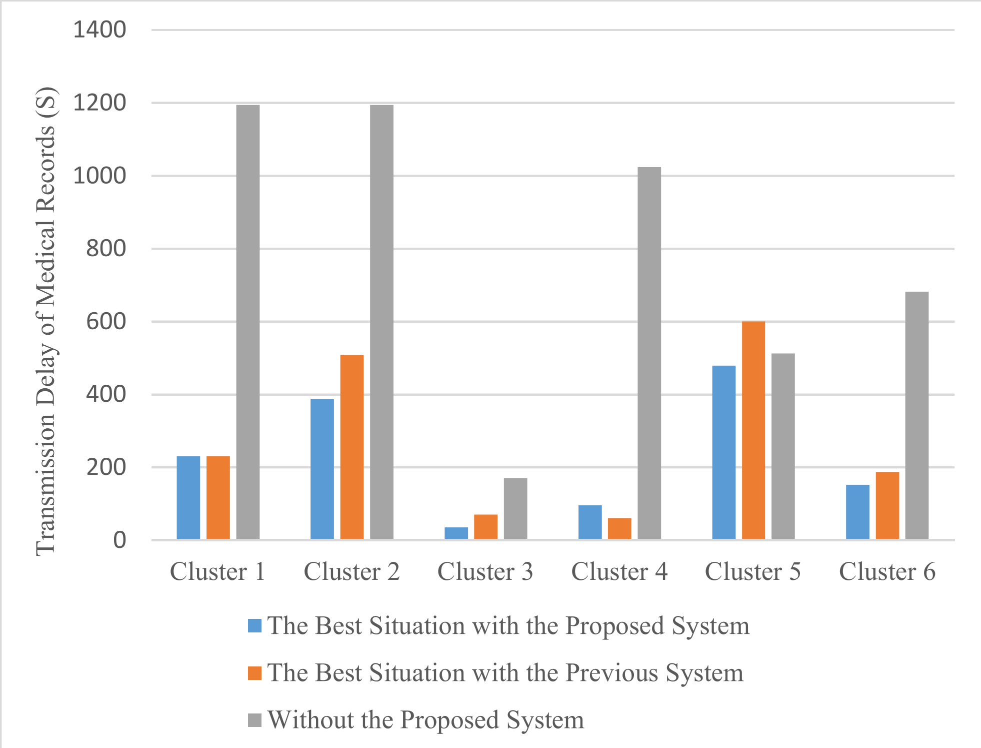

Figure 5.7: The Transmission Delay Results of the Proposed System Based on Deep Q- Learning.

As we can find from the table, it costs 476.97s for all patients finished their caching process in proposed system. However, in our last system, it costs 600.51s for all the patients finish their caching, which saves 20.57% time for other patients who want to cache in these three edge devices. The total transmission time without the proposed system is 1194.67s, which is a 60.08% improvement with new proposed system.

The system improvement of average transmission delay is 71.14% comparing with the scenario without using deep learning proposed system. However, by applying the previous system, the improvement of average transmission delay is 65.32%, which is more than 16.78% improvement.

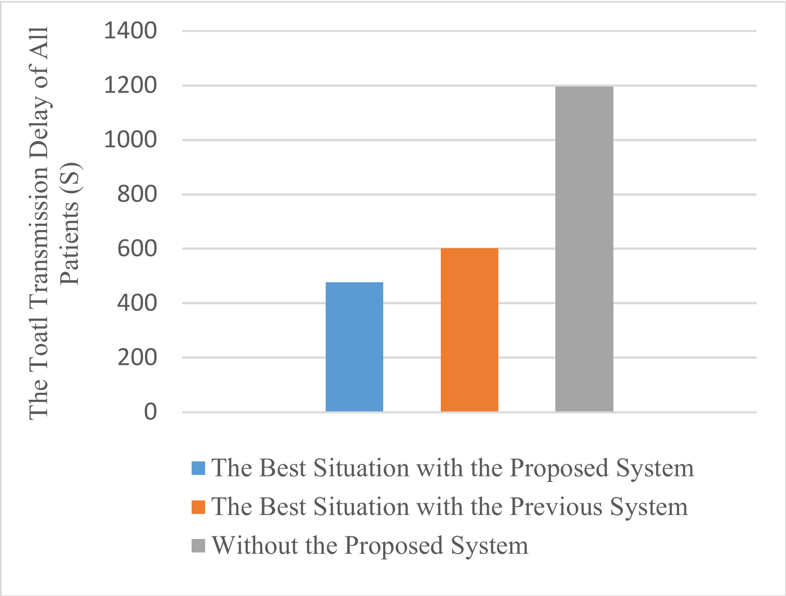

Figure 5.8: The Total Transmission Time for All the Patients.

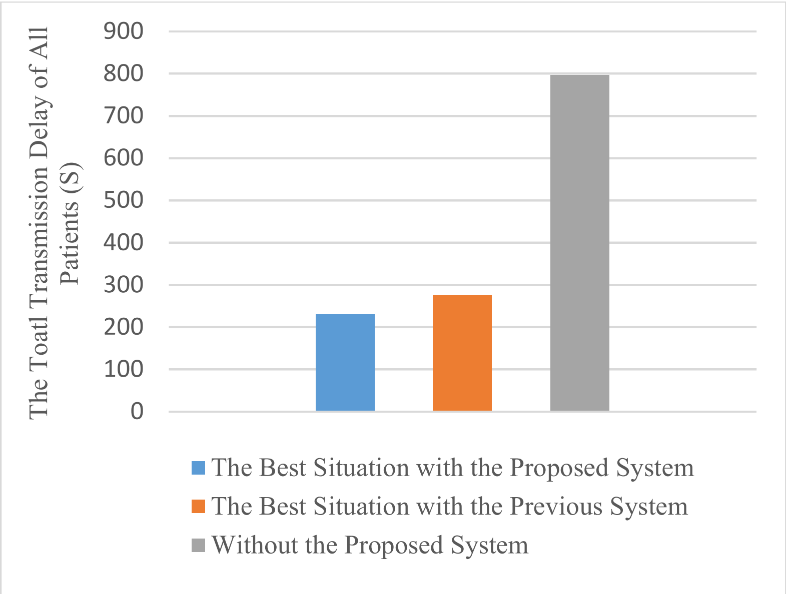

Figure 5.9: The Average Transmission Delay of Three Scenarios.

5.4 Summary and Conclusions

In this research, we introduced a new allocation algorithm based on deep Q-learning algorithm for patients to cache their medical files to adjacent edge devices. In the proposed system, the medical records can be cached in multiple edge caching devices dispersedly at same time instead of one edge caching only. The processing time including queueing time and caching time is decreased significantly. The learning procedure of the system is to find the optimal Q value for each patient. Q-learning is proposed for patients to find the best routine to cache their medical records among the edge devices. The new system provides more fair and humanized medical care services comparing with previous system based on game theory. The patients with higher level of illness severity have higher priority and lower waiting time to be served. According to the simulation results, patients could get more efficient medical care by applying the proposed system and save at least 71.14% average transmission time comparing the situation of traditional method. We compared total transmission time as well, the results show that new allocation algorithm saves 20.57% and 60.08% time for the whole of system comparing with the previous system and the scenario without proposed algorithm, which means other patients who want to cache in these three edge devices could get medical care in a quicker manner in our new system.

Chapter 6
Conclusions and Future Work

6.1 Conclusions

In this research, novel and advanced telecommunication technologies have been proposed to access medical records with low transmission delay and save more time for clinical professions and patients. The contributions of this thesis cover the simulation and performance analysis of three different scenarios, including the first system with location-aware femtocaching and DVS camera, the second system with edge computing and auction-based non-cooperated game theory, and the third system with deep Q-learning. The main conclusions of this thesis are summarized as follows.

In the first system, we allocate patient's medical records to the nearest hospitals of home area, work place, family home, friends' home and other places according to the patient's social life. Location-aware femtocaching has been used to cache patients' EMR and set up close to all the hospitals. It is the first time that femtocaching technology is proposed in healthcare area. We divided electronic medical records into three levels according to the type and the size of medical files (word/text, image, and video/audio). The EMR is allocated in femtocaches efficiently and effectively according to our proposed medical records allocation algorithm by applying knapsack model with penalty minimization. Due to the limited storage capacity of caches, not all the caches have enough space for medical files, especially for the medical images and videos. The proposed medical records allocation algorithm would choose the most suitable files to cache based on three important penalty parameters including the staying time of patients in one location, the importance of medical files to patient's current health condition, and the transmission delay of medical files. The purpose is to provide medical files in a timely manner and reduce delay when medical files are being shared and allocated among the hospitals. The simulation results show that the proposed scheme can drop the transmission delay of EMR by 43.57% to 86.66% compared with the scenario without using location-aware femtocaching technology. In the proposed system, the medical video accounts for a big part of the EMR (around 70%), which causes a huge delay when physicians access to patient's medical records. Medical video is a significant part of health care area. Video was selected not only because it provides important reference to physicians, but also providing an additional option for close supervision and offered increased privacy and as a result, decreasing stress for the patients. A new medical video capture method is proposed to reduce the size of visual medical files by applying DVS camera technology. The DVS camera only takes video when patient is moving. DVS camera could improve the bandwidth utilization and save more resources, such as battery life of devices and storage capacity of camera to current monitoring system, especially for the monitoring of patients. The transmission delay in proposed system can be decreased significantly when transferring or sharing the medical files among the hospitals. In particular, for the first time, exploiting the recently developed DVS technology is investigated here to achieve smaller medical visual data files in our proposed tele-health system. The simulation results show that the proposed scheme by adding DVS camera technology can decrease the transmission delay by 41% to 80.4% more comparing with femtocaching only, which means the transmission delay can be improved by 88.94% to 92.13%.

In the previous sections, the femtocaching is located to the nearest hospital where patients stay. Due to the changing location of patients, it is difficult to allocate and cache medical records to all the nearest hospital. Once emergency happens in other places, the physician cannot access medical records of patients and provide timely healthcare services. Edge computing is proposed to process and cache EMR, which is an advanced telecommunication technology that users can process their data on the edge before sending to the public cloud. Comparing with traditional cloud computing, Edge computing could off load partly or whole data processing from cloud and provide more efficient and safer services. Patient can control their own medical data, made by wearable sensors, on the edge. On the other hand, in the last chapter, patients only can use their own femtocaching to cache and allocate EMR. However, in the proposed system, patients can share their edge caching with others. A new medical files allocation algorithm has been applied by using action-based and non-cooperative game theory. Four medical factors including the severity of illness, the size of the medical records that patient wants to cache, the queueing time that patients have waited, and the rewards that patients got by sharing edge caching with other patients, have been combined with telecommunication channel capacity for the patients to compete the allocation and caching priority. Resolving the competition for finite storage capacity among other patients, we model the patient's action as a game, which requests the constrained storage capacity from the edge devices. To achieve the optimal results, we formulated the storage request strategies as auction-based strategies and developed the corresponding non-cooperate game mechanisms. We assume that patients who request storage are selfish, and each patient rationally behaves to maximize its own benefit. Patients use weights as a bid for storage capacity and each edge device may assign the extra storage among the patients by itself according to the amount of medical points and channel condition. The patients calculate their bid with proposed game mechanism and send to edge devices. After several rounds' competition, the edge devices would choose the winner and cache the medical files by order. We analyzed the scenarios with 6, 60, 600, 6000 and 60,000 patients. The simulation results show that the proposed system can decrease the transmission delay by 56.87% to 93.69% in the best situation (the cached medical records are enough for medical support) comparing the system without sharing edge device, which means the proposed system saves more than half time for clinical professors to provide efficient medical services. In the worst situation (the cached medical records are not enough for medical support and whole EMR is required), the simulation results show that the proposed system can decrease the transmission delay by 25.94% to 57.75%. The results also show that the more shared edge devices, the higher channel capacity patients can get, which means the proposed system encourages more patients to share their storage capacity of edge devices to others.

In our last system, we considered that all the medical factors have same weight. However, the priority is based on a combination value. It is possible that for the patients who has higher illness servility but cannot get timely and effective health care services. On the other hand, in the last system, the medical files of patients can only be cached in a single edge device which could provide the fast services. However, in our new system, medical records can be cached in multiple edge caching devices dispersedly at same time. The processing time including queueing time and caching time would be decrease significantly. A new sharing algorithm based on deep Q-learning is proposed to allocate and cache the medical records of multiple patients. Four states are considered in the allocation algorithm including the illness servility of patients, the storage capacity of edge devices, the processing time of medical files and the channel condition between patients and edge devices. The action of this learning algorithm is the caching decision (1/0 or yes/no) of edge devices for each patient. The learning procedure of the system is to find the optimal Q value for each patient. Deep Q-learning is proposed for patients to find the best routine to cache their medical records among the edge devices. M/M/1 queueing model with preemptive priority is proposed for queueing time of patients. The new system provides more fair and humanized medical care services comparing with the last one. The patients with higher level of illness severity have higher priority and lower waiting time to be served. According to the simulation results, patients could get more efficient medical care by applying the proposed system and save at least 16.78% and 71.14% average transmission time comparing the last system and the traditional method. We compared total transmission time as well, the results show that new allocation algorithm could save 20.57% and 60.08% time for the whole of system comparing with the previous system and the scenario without proposed algorithm, which means other patients who want to cache could get medical care in a quicker manner in our new system.

6.2 Future Works

The thesis covers femtocaching, edge computing, DVS camera, auction-based non-cooperated game theory and deep Q-learning in health care networks and provides reference for future research. Several challenges that need to be addressed remain in this research. The first one is the security of medical files that cached on other patients' edge devices. A strict privacy mechanism is needed. Secondly, for the part of channel condition, the scenario we considered is with an ideal environment (white noise only). However, when more patients share their edge devices, the channel interference among the patients should be considered. Thirdly, a fairness mechanism is needed for patient to ensure every patient get efficient and effective medical services.

Reference

[1] D. A. Perednia, A. Allen, "Telemedicine Technology and Clinical Applications", *JAMA,* vol. 273(6), pp. 483-488, 1995.

[2] A. W. Templeton, S. J. Dwyer III, J. A. Johnson, W. H. Anderson, K. S. Hensley, K. R. Lee, S. J. Rosenthal, D. F. Preston, and S. Batnitzky, "Implementation of an on-line and long term digital management system", *Radiographics*, vol. 5, no. 1, pp. 121-138, 1985.

[3] A. Shachak, S. Reis, "The impact of electronic medical records on patient–doctor communication during consultation: a narrative literature review", *Journal of Evaluation in Clinical Practice*, vol. 15, no. 4, pp. 641-649, 2009.

[4] A. C. Smith, M. Bensink, N. Armfield, J. Stillman, L. Caffery, "Telemedicine and rural health care applications", *Journal of Postgraduate Medicine,* vol. 51, no. 4, pp. 286-293, 2005.

[5] B. Peter, M. D. Angood, "Telemedicine, the Internet, and World Wide Web: Overview, Current Status, and Relevance to Surgeons", *World Journal of Surgery,* vol. 25, no. 11, pp. 1449-1457, Nov. 2001.

[6] H. Lehrach, A. Ionescu, "The Future of Health Care: deep data, smart sensors, virtual patients and the Internet-of-Humans", *Digital4Science,* 2016.

[7] E. A. Miller, "Telemedicine and the Provider-Patient Relationship: What We Know So Far", *The Ethics of 'Personalised' Medicine in a Consumer Age,* Jan. 2010.

[8] P. Whitten, B. Love, "Patient and provider satisfaction with the use of telemedicine: Overview and rationale for cautious enthusiasm", *J Postgrad Med*, vol. 51, pp. 294-300, 2005.

[9] S. J. Wang, B. Middleton, L. A. Prosser, C. G. Bardon, C. D. Spurr, P. J. Carchidi, A. F. Kittler, R. C. Goldszer, D. G. Fairchild, A. J. Sussman, G.J. Kuperman, D. W. Bates, “A cost-benefit analysis of electronic medical records in primary care”, *The American Journal of Medicine*, Vol. 114, no. 5, pp. 397-403, Apr. 2003.

[10] C. Liu, A. Long, Y. Li, K. Tsai, H. Kuo, “Sharing patient care records over the World Wide Web”, *National Library of Medicine*, vol. 61(2-3), pp. 189-205, 2001.

[11] M. D. Rodríguez, J. Favela, E. A. Martínez, M. A. Muñoz, “Location-Aware Access to Hospital Information and Services,” IEEE Transactions on information technology in biomedicine, Vol. 8, No. 4, 2004.

[12] C. F. Snyder, A. W. Wu, R. S. Miller, R. E. Jensen, E. T. Bantug, A. C. Wolff, “The Role of Informatics in Promoting Patient-Centered Care”, *The Cancer Journal*, vol. 17(4), pp. 211-218, Jul. 2011.

[13] F. Sarhan, “Telemedicine in healthcare: exploring its uses, benefits and disadvantages”, *Nursing Times.net*, 2009.

[14] M. Jonathan, B. Gerald-Mark, “Telemedicine: Its Effects on Health Communication”, *Health Communication*, vol. 21 (1), pp. 73–83, 2007.

[15] B. Dinesen, B. Nonnecke, D. Lindeman, et al, “Personalized Telehealth in the Future: A Global Research Agenda”, *J Med Internet Res*, 18(3):e53, Mar. 2016.

[16] H. Fraser, P. Biondich, D. Moodley, S. Choi, B. Mamlin, P. Szolovits, “Implementing electronic medical record systems in developing countries”, *Informatics in Primary Care*, vol. 13, pp. 83–95, 2005.

[17] N. Golrezaei, A. F. Molisch, A. G. Dimakis, G. Caire, “Femtocaching and device-to- device collaboration: A new architecture for wireless video distribution,” *Communications Magazine, IEEE,* vol. 51, no. 4, pp. 142-149, 2013.

[18] N. Golrezaei, K. Shanmugam, A. G. Dimakis, A. F. Molisch, G. Caire, “Femtocaching: Wireless video content delivery through distributed caching helpers,” *INFOCOM, IEEE,* pp. 1107-1115, 2012.

[19] M. Litzenberger, C. Posch, D. Bauer, A.N. Belbachir1, P. Schön, B. Kohn, H. Garn, “Embedded Vision System for Real-Time Object Tracking using an Asynchronous Transient Vision Sensor”, *12th Digital Signal Processing Workshop & 4th IEEE Signal Processing Education Workshop, IEEE,* pp. 173-178, Sept, 2006.

[20] C. Posch, R. Benosman, R.E. Cummings, “Giving Machines Humanlike-Viosn similar to our own would let devices capture images more efficiently”, *Spectrum. IEEE*, Dec, 2015.

[21] E. Grenet, S. Gyger, P. Heim, F. Heitger, F. Kaess, P. Nussbaum, P. F. Ruedi, “High Dynamic Range Vision Sensor for Automotive Applications”, *SPIE,* vol. 5663, pp. 246-253, 2005.

[22] P. Lichtsteiner, C. Posch, T. Delbruck, “A 128 128 120 dB 15 us Latency Asynchronous Temporal Contrast Vision Sensor”, *IEEE Journal of solid-state circuits*, vol. 43, no. 2, Feb, 2008.

[23] C. O. Rolim, F. L. Koch, C. B. Westphall, J. Werner, A. Fracalossi, G. S. Salvador, “A Cloud Computing Solution for Patient's Data Collection in Health Care Institutions”, *Second International Conference on eHealth, Telemedicine, and Social Medicine*, 2010.

[24] M. Li, S. Yu, Y. Zheng, K. Ren, W. Lou, “Scalable and Secure Sharing of Personal Health Records in Cloud Computing Using Attribute-Based Encryption”, *IEEE Transactions on parallel and distributed systems*, vol. 24, no. 1, pp. 131-143, 2012.

[25] W. Shi, J. Cao, Q. Zhang, Y. Li, L. Xu, “Edge computing: Vision and Challenges”, 2015.

[26] S. Oulaourf, A. Haidine, H. Ouahmane, “Review on using Game Theory in Resource Allocation for LTE/LTE-Advanced,” *2016 International*

Conference on Advanced Communication Systems and Information Security (ACOSIS), pp. 1-7, Feb. 2017.

[27] A. Nezarat, G. Dastghaibifard, "Efficient Nash Equilibrium Resource Allocation Based on Game Theory Mechanism in Cloud Computing by Using Auction," *PLoS ONE,* vol. 10, no. 10, Oct. 2015.

[28] X. Zhang, Q. Zhu, "Game-Theory Based Power and Spectrum Virtualization for Optimizing Spectrum Efficiency in Mobile Cloud-Computing Wireless Networks," *IEEE Transactions on Cloud Computing,* pp. 1-1, Jul. 2017.

[29] J. Schmidhuber, "Deep learning in neural networks: An overview", *Neural Networks,* vol. 61, Pages 85-117, Jan. 2015.

[30] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, D. Wierstra, "Continuous Control with Deep Reinforcement Learning", *ICLR,* 2016.

[31] D. G. Páez, F. Aparicio, M. de Buenaga , J. R. Ascanio, " Chronic Patients Monitoring Using Wireless Sensors and Big Data Processing", *Eighth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing December*, 2014.

[32] G. Appelboom, E. Camacho, M. E. Abraham, etc., "Smart wearable body sensors for patient self-assessment and monitoring", *Archives of Public Health*, Aug. 2014.

[33] F. Fernandez, G. C. Pallis, "Opportunities and challenges of the Internet of Things for healthcare", *4th International Conference on Wireless Mobile Communication and Healthcare - Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH)*, Jan. 2015.

[34] S. Patel, H. Park, P. Bonato, L. Chan, M. Rodgers, "A review of wearable sensors and systems with application in rehabilitation", *Journal of Neuro Engineering and Rehabilitation* 2012.

[35] P. Corbishley, E. Rodriguez-Villegas, "Breathing Detection: Towards a Miniaturized, Wearable, Battery-Operated Monitoring System", *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 1, pp. 196-204, Jan. 2008.

[36] M. Lanz, A. Nahapetianz, A. Vahdatpourz, L. Kaiserx, M. Sarrafzadeh, "SmartFall: an automatic fall detection system based on subsequence matching for the SmartCare", *International Conference on Body Area Networks*, 2009.

[37] J. Patterson, D. G. McIlwraith, Y. Guang-Zhong, "A Flexible, Low Noise Reflective PPG Sensor Platform for Ear-Worn Heart Rate Monitoring", *Wearable and Implantable Body Sensor Networks, BSN Sixth International Workshop*, pp. 286-291, Jun. 2009.

[38] C. Perera, A. Zaslavsky, P. Christen, D. Georgakopoulos, "Context Aware Computing for The Internet of Things: A Survey", *IEEE Communications Surveys & Tutorials,* vol. 16, no. 1, pp. 414 – 454, 2014.

[39] H. Sundmaeker, P. Guillemin, P. Friess, and S. Woelffle, "Vision and challenges for realising the internet of things", *European Commission Information Society and Media*, Mar. 2010.

[40] A. Zaslavsky, C. Perera, and D. Georgakopoulos, "Sensing as a service and big data," *International Conference on Advances in Cloud Computing, ACC*, pp. 21–29, Jul. 2012.

[41] H. Chen, S. Compton, O. Hsiao, "DiabeticLink: A Health Big Data System for Patient Empowerment and Personalized Healthcare", *International Conference on Smart Health*, pp. 71-83, 2013.

[42] G. Yang, L. Xie, M. Mäntysalo, X. Zhou, Z. Pang, L. Xu, S. Kao-Walter, Q. Chen, L. Zheng, "A Health-IoT Platform Based on the Integration of Intelligent Packaging, Unobtrusive Bio-Sensor, and Intelligent Medicine Box", *IEEE Transaction on Industrial Informatics*, vol. 10, no. 4, Nov. 2014.

[43] P. P. Ray, "Home Health Hub Internet of Things (H3 IoT): An Architectural Framework for Monitoring Health of Elderly People", *International Conference on Science Engineering and Management Research (ICSEMR)*, Feb. 2015.

[44] B. B. Dean, J. Lam, J. L. Natoli, Q. Butler, D. Aguilar, and R. J. Nordyke, "Review: Use of Electronic Medical Records for Health Outcomes Research," *Medical Care Research and Review*, vol. 66, no. 6, pp. 611-638, Dec. 2009.

[45] P. J. O'Connor, A. L. Crain, W. A. Rush, J. M. Sperl-Hillen, J. J. Gutenkauf, J. E. Duncan, "Impact of an Electronic Medical Record on Diabetes Quality of Care," *Annals of Family Medicine*, vol. 3, no. 4, pp. 300-6, Jul. 2005.

[46] K. Spaziano, "Electronic Medical Records," *Radiologic Technology,* vol. 72, no. 3, pp. 287, Jan. 2001.

[47] K. D. Mandl, P. Szolovits, I. S. Kohane, "Public Standards and Patients' Control: How to Keep Electronic Medical Records Accessible but Private," *BMJ (Clinical research ed.),* vol. 322, no. 7281, pp.283-7, Feb. 2001.

[48] E. Y. S. Lim, M. Fulham, D. D. Feng, "Biomedical Information Technology," *Biomedical Engineering*, pp. 29-49, 2008.

[49] G. Perera, A. Holbrook, L. Thabane, G. Foster, D. J. Willison, "Views on Health Information Sharing and Privacy From Primary Care Practices Using Electronic Medical Records," *International Journal of Medical Informatics*, vol. 80, no. 2, pp. 94-101, Feb. 2011.

[50] G. Eysenbach, A. R. Jadad, "Evidence-based Patient Choice and Consumer health informatics in the Internet age", *Journal of medical internet research*, Jun. 2001.

[51] A. Pingyod, Y. Somchit, "Content Updating Method in FemtoCaching", *11th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, Jun. 2014.

[52] J. Won, H. Ryu, T. Delbruck, J. H. Lee, J. Hu, "Proximity Sensing Based on a Dynamic Vision Sensor for Mobile Devices", *IEEE Transactions on Industrial Electronics*, vol. 62, no. 1, Jan. 2015.

[53] A. Ahmed, E. Ahmed, "A Survey on Mobile Edge Computing", *10th International Conference on Intelligent Systems and Control (ISCO),* Nov. 2016.

[54] M. T. Beck, M. Werner, S. Feld, T. Schimper, "Mobile Edge Computing: A Taxonomy", *The Sixth International Conference on Advances in Future Internet, AFIN*, 2014.

[55] M. Satyanarayanan, P. Bahl, R. Caceres, N. Davies, "The case for vm-based cloudlets in mobile computing," *Pervasive Computing, IEEE*, vol. 8, no. 4, pp. 14–23, 2009.

[56] K. Gai, M. Qiu, H. Zhao, L. Tao, Z. Zong, "Dynamic energy-aware cloudlet-based mobile cloud computing model for green computing", *Journal of Network and Computer Applications*, vol 59, pp. 46-54, Jan 2016.

[57] D. T. Hoang, D. Niyato. P. Wang, "Optimal admission control policy for mobile cloud computing hotspot with cloudlet", *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 3145-3149, 2012.

[58] Y. Jararweh, L. Tawalbeh, F. Ababneh, F. Dosari, "Resource Efficient Mobile Computing Using Cloudlet Infrastructure", *IEEE 9th International Conference on Mobile Ad-hoc and Sensor Networks,* pp. 373-377, 2013.

[59] K. Habak, M. Ammar, K. Harras, E. Zegura, "Femto clouds: Leveraging mobile devices to provide cloud service at the edge", *IEEE 8th International Conference on Cloud Computing (CLOUD)*, pp. 9–16, June 2015.

[60] S. Abdelwahab, B. Hamdaoui, M. Guizani, T. Znati, "Replisom : Disciplined tiny memory replication for massive iot devices in lte edge cloud", *Internet of Things Journal, IEEE*, vol. PP, no. 99, pp. 1–1, 2015.

[61] R. Trestian, O. Ormond, G. Muntean, "Game Theory-Based Network Selection: Solutions and Challenges", *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, 2012.

[62] Y. LeCun, Y. Bengio, G. Hinton, "Deep learning", *Nature*, 2015.

[63] R. Miotto, F. Wang, S. Wang, X. Jiang, J. T. Dudley, "Deep learning for healthcare: review, opportunities and challenges", *Briefings in Bioinformatics*, pp. 1-11, 2017.

[64] Y. Lecun, L. Bottou, Y. Bengio, et al., "Gradient-based learning applied to document recognition", *Proc IEEE*, 1998.

[65] R.J. Williams, D. Zipser, "A learning algorithm for continually running fully recurrent neural networks", *Neural Comput*, 1989.

[66] C. Liu, F. Wang, J. Hu, et al., "Risk prediction with electronic health records: a deep learning approach", *ACM International Conference on Knowledge Discovery and Data Mining,* pp. 705-714, 2015.

[67] T. Pham, T. Tran, D. Phung, et al., "DeepCare: a deep dynamic memory model for predictive medicine", arXiv, 2016, https://arxiv.org/abs/1602.00357.

[68] E. Choi, M. Bahadori, A. Schuetz, et al., "Doctor AI: predicting clinical events via recurrent neural networks," ArXiv, 2015, http://arxiv.org/abs/1511.05942v11.

[69] S. Pakhomov, S. J. Jacobsen, C. G. Chute, V. L. Roger, "Agreement between Patient-reported Symptoms and their Documentation in the Medical Record", *US National Library of Medicine*, 2008.

[70] B. B. Dean, J. Lam, J. L. Natoli, Q. Butler, D. Aguilar, R. J. Nordyke, "Review: Use of Electronic Medical Records for Health Outcomes Research", *Medical Care Research and Review,* vol. 66, no. 6, pp. 611-638, Dec. 2009.

[71] P. J. O'Connor, A. L. Crain, W. A. Rush, J. M. Sperl-Hillen, J. J. Gutenkauf, J. E. Duncan, "Impact of an Electronic Medical Record on Diabetes Quality of Care", *Annals of Family Medicine*, vol. 3, no. 4, pp. 300-6, Jul. 2005.

[72] I. Iakovidis, "Towards personal health record: current situation, obstacles and trends in implementation of electronic healthcare record in Europe", *International Journal of Medical Informatics,* vol 52, Issues 1–3, pp. 105-115, 1998.

[73] Z. Chen, M. Shikh-Bahaei, "Location-Aware Distributed File Allocation with Femtocaching and DVS for Low-Delay Access to Electronic Medical Records", *IEEE Engineering in Medicine and Biology Society, EMBC*, 2016.

[74] J. Xiang, Y. Zhang, T. Skeie, L. Xie, "Downlink Spectrum Sharing for Cognitive Radio Femtocell Networks", *IEEE Systems Journal*, vol. 4, no. 4, pp. 524-534, Dec. 2010.

[75] P. Kulkarni, W. H. Chin, T. Farnham, "Radio resource management considerations for lte femto cells", *ACM SIGCOMM Comput. Commun. Rev.*, vol. 40, no. 1, pp. 26–30, 2010.

[76] S. Lien, C. Tseng, K. Chen, C. Su, "Cognitive Radio Resource Management for QoS Guarantees in Autonomous Femtocell Networks", *2010 IEEE International Conference on Communications*, pp. 1-6, 2010.

[77] E. Kaasinen, "User needs for location-aware mobile services", *Personal and Ubiquitous Computing*, vol. 7, Issue. 1, pp. 70–79, May. 2003.

[78] W. Leung, M. Shikh-Bahaei. "A New Femtocaching File Placement Algorithm for Telemedicine", Proceedings, *IEEE Engineering in Medicine and Biology Society, EMBC 2015,* 2015.

[79] Galbraith J, et al, "Cost analysis of a falls-prevention program in an orthopedic setting", *Clinical Orthopedics and Related Research,* 2011

[80] B. Mwenge, A. Brion, G. Uguccioni, I. Arnulf, "Sleepwalking: long-term home video monitoring", *Sleep Medicine, Elsevier,* pp. 1226 – 1228, 2013.

[81] J. Davis1 , M. Kutash, J. Whyte IV, "A comparative study of patient sitters with video monitoring versus in-room sitters", *Journal of Nursing Education and Practice*, vol. 7, no. 3, 2017.

[82] J. Won, H. Ryu, T. Delbruck, J. H. Lee, J. Hu, "Proximity Sensing Based on a Dynamic Vision Sensor for Mobile Devices", *IEEE Transactions on industrial electronics*, vol. 62, no. 1, Jan. 2005.

[83] S. Ajami, T. Bagheri-Tadi, "Barriers for Adopting Electronic Health Records (EHRs) by Physicians", *ACTA INFORMATICA MEDICA*, vol. 21, no. 2, pp. 129-134, 2013.

[84] S. Patel, H. Park, P. Bonato, L. Chan, M. Rodgers, "A Review of Wearable Sensors and Systems with Application in Rehabilitation", *Journal of neuroengineering and rehabilitation,*2012.

[85] Case Study 4: Sleep disorder research, http://inilabs.com/products/dynamic-vision-sensors/dvs-case-studies/.

[86] Z. Chen, T. Shikh-Bahaei, P. Luff, M. Shikh-Bahaei, "Edge caching and Dynamic Vision Sensing for low delay access to visual medical information", *IEEE Engineering in Medicine and Biology Society, EMBC,* 2017.

[87] R. M. Freund, "Penalty and Barrier Methods for Constrained Optimization", *Massachusetts Institute of Technology*, 2004.

[88] E. R. Bogoch, V. Elliot-Gibson, D. E. Beaton, S. Jamal, R. Josse, T. M. Murray, "Effective Initiation of Osteoporosis Diagnosis and Treatment for Patients with a Fragility Fracture in an Orthopedic Environment," *J Bone Joint Surg Am*, 2006.

[89] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, "Wireless sensor networks: a survey Computer Networks", *ELSEVIER*, vol. 38, Issue. 4, pp. 393-422, Mar. 2002.

[90] J. Kingman, L. Rogers, P. Baxendale, P. Greenwood, F. Kelly, J. Legall, E. Pardoux, D. Williams, “Oxford studies in probability: Poisson Processes”, *Oxford University Press,* 2002.

[91] Q. Zhang, L. Cheng, R. Boutaba, “Cloud computing: state-of-the-art and research challenges”, *Journal of Internet Services and Applications*, Vol. 1, Issue. 1, pp. 7 – 18, May. 2010.

[92] S. Marstona, Z. Lia, S. Bandyopadhyaya, J. Zhanga, A. Ghalsasib, “Cloud computing -the business perspective”, *Decision Support System, Elsevier,* Vol. 51, No. 1, pp. 176-189, 2011.

[93] M. Satyanarayanan, P. Simoens, Y. Xiao, P. Pillai, Z. Chen, K. Ha, W. Hu, B. Amos, “Edge Analytics in the Internet of Things”, *IEEE Pervasive Computing*, Vol. 14, Issue. 2, pp. 24-31. 2015.

[94] S. Abdelwahab, B. Hamdaoui, M. Guizani, T. Znati, “Replisom : Disciplined tiny memory replication for massive iot devices in lte edge cloud,” *Internet of Things Journal, IEEE,* vol. PP, no. 99, pp. 1–1, 2015.

[95] A. B. MacKenzie, S. B. Wicker, “Game theory in communications: motivation, explanation, and application to power control”, *IEEE Global Telecommunications Conference,* Aug. 2002.

[96] A. B. MacKenzie, S. B. Wicker, “Game theory and the design of self-configuring, adaptive wireless networks”, *IEEE Communications Magazine*, vol. 39, no. 11, pp. 126-131, Nov. 2001.

[97] D. E. Charilas, A. D. Panagopoulos, “A Survey on Game Theory Applications in Wireless Networks”, *Computer Networks*, *Elsevier*, Vol. 54, pp. 3421-3430. 2010.

[98] Z. Han, Z. Ji, K. J. R. Liu, “Power Minimization for Multi-Cell OFDM Networks Using Distributed Non-Cooperative Game Approach”, *Globecom, IEEE,* 2004.

[99] Y. Xiao, X. Shan, Y. Ren, “Game Theory Models for IEEE 802.11 DCF in Wireless Ad Hoc Networks”, *IEEE Radio* Communications, Mar. 2005.

[100] X. Wang, Z. Li, P. Xu, Y. Xu, X. Gao, H. Chen, “Spectrum Sharing in Cognitive Radio Networks—An Auction-Based Approach”, *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* , vol. 40, no. 3, pp. 587-596, Jun. 2010.

[101] J. Jia, Q. Zhang, Q. Zhang, M. Liu, “Revenue Generation for Truthful Spectrum Auction in Dynamic Spectrum Access”, *The Tenth ACM International Symposium on Mobile ad Hoc Networking and Computing*, pp. 3-12, May. 2009.

[102] A. Kakhbod, A. Nayyar, D. Teneketzis, “Revenue Maximization in Spectrum Auction for Dynamic Spectrum Access”, *The 5th International ICST Conference on Performance Evaluation Methodologies and Tools*, pp. 558-564, May. 2011.

[103] J. Huang, R. A. Berry, M. L. Honig, “Auction-Based Spectrum Sharing”, *Mobile Networks and Applications,* vol. 11, pp. 405-418, Apr. 2006.

[104] L. Makowski, J. M. Ostroy, “Vickrey-Clarke-Groves mechanisms and perfect competition”, *Journal of Economic Theory,* vol. 42, no. 2, pp. 244-261, Aug. 1987

[105] P. Chaikijwatana, T. Tachibana, “VCG Auction-based Bandwidth Allocation with Network Coding in Wireless Networks”, *Applied Computer and Applied Computational Science*, 2011.

[106] A. Jarray and A. Karmouch, "VCG auction-based approach for efficient Virtual Network embedding," 2013 IFIP/IEEE International Symposium on Integrated Network Management (IM 2013), Ghent, 2013, pp. 609-615.

[107] S. D. Horn, R. A. Horn, P. D. Sharkey, “The Severity of Illness Index as A Severity Adjustment to Diagnosis-related Groups”, *Heath Care Financing Review*, Nov. 1984.

[108] L. Woodworth, P. S. Romano, J. F. Holmes, “Does Insurance Status Influence a Patient’s Hospital Charge?”, *Appl Health Econ Health Policy*, vol. 15, no. 3, pp. 353-362, Jun. 2017.

[109] S. Verdu, T. S. Han, “A general formula for channel capacity”, *IEEE Transactions on Information Theory,* vol. 40, no. 4, Jul. 1994.

[110] K. Kafadar, “Gaussian white-noise generation for digital signal synthesis”, *IEEE Transactions on Instrumentation and Measurement*, vol. IM-35, Issue. 4, Dec. 1986.

[111] S. Gu, T. Lillicrap, I. Sutskever, S. Levine, “Continuous Deep Q-Learning with Model-based Acceleration”, *Proceedings of the 33rd International Conference on Machine Learning*, volume 48, 2016

[112] V. Mnih, et al., “Human-level control through deep reinforcement learning”, *Nature*, 2015.

[113] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, “Human-level control through deep reinforcement learning”, *Nature 518*, pp.529-533, 2015.

[114] M. K. Tarabia, “Analysis of M/M/1 Queueing System with Two Priority Classes”, *OPSEARCH*, vol. 44, no.4, 2007.